ANALYSIS OF SMALL MUSCLE MOVEMENT EFFECTS ON EEG SIGNALS

THESIS

Erhan Erdi YANTERI, First Lieutenant, TuAF

AFIT-ENG-MS-16-D-051

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENG-MS-16-D-051

ANALYSIS OF SMALL MUSCLE MOVEMENT EFFECTS ON EEG SIGNALS

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science (Computer Engineering)

Erhan E YANTERI

First Lieutenant, TuAF

December 2016

AFIT-ENG-MS-16-D-051

ANALYSIS OF SMALL MUSCLE MOVEMENT EFFECTS ON EEG SIGNALS

Erhan E YANTERI

First Lieutenant, TuAF

Committee Membership:

Dr. Brett J. Borghetti
Chair

Dr. Michael E. Miller
Member

Mr. Justin R. Estepp
Member

AFIT-ENG-MS-16-D-051

Abstract

Developments in the biomedical signal processing have led the electroencephalography (EEG) to be a critical tool for the Brain Computer Interface (BCI) systems and Human Machine Teams (HMTs). Both of them strongly rely on the EEG signals in order to evaluate the neural activity and the cognitive state. They need to use the pure EEG signal that only represents the neural activity of the brain, but the physiological and non-physiological artifacts distort the EEG signal and make the interpretation of cognitive state harder or they may cause misinterpretations.

While developing teams of humans and computer agents, certain human activities are essential. While interacting with computers, humans perform small motor muscle movements such as operating a keyboard and mouse, manipulating a stick and throttle, or performing touch-screen activities. On the other side, the computer agent needs to know the cognitive state of the human teammate in order to make decisions and the EEG signals are the only information source of cognitive state.

In this thesis, the artefactual effects of the small muscle movements such as hand and finger movements that are necessary for the keyboard and mouse usage are investigated. Five males (all right handed) participate in this study and there are two sessions in different days for each participant. Six different conditions are recorded in this experiment. These conditions are the resting state, left finger keyboard press, right finger keyboard press, imaginary right finger press, mouse movement and the activity condition that includes both the finger presses and the mouse movements.

The EEG data is recorded from 9 channels with a sampling rate of 2000 Hz. Two channels EMG data from the right arm, two channel EOG data and the digitals channels that indicate the keyboard and mouse presses are collected in addition to the EEG signal.

Upper frequency bands (>30 Hz) of the EEG signal are extracted in order to investigate the artefactual effects of the small muscle movements. First, the activity and resting conditions are compared in order to investigate the effects of the small muscle movements when the contamination level is high. Time and frequency domain features are extracted from the upper frequency bands and some time domain features such as variance obtain a visual difference between two states. Three-layer neural network model is created for the classification of the activity and resting states and the model yields 92.2% accuracy.

Secondly, the right and left finger press conditions are investigated for the finger artifacts. Analyzing of the time and frequency domain features shows that the effects of the finger artifacts are not visually observable. But the neural network model obtains 64% classification accuracy for the finger artifact detection. The classification accuracy of the finger artifacts increases to 72% after removing the eye blink artifacts.

The results of the classification support our hypothesis about the artefactual effects of the small muscle movements.

Acknowledgment

I would like to express my most sincere appreciation to my thesis advisor, Dr. Brett J. Borghetti. I always felt his support with me. Without his wisdom, guidance and patience, I could not complete this thesis study. He was the one motivating and encouraging me throughout this challenging journey.

I would like to thank to Dr. Michael E. Miller and Mr. Justin R. Estepp for accepting to be in my committee. Their insights and suggestions have made great contributions to my research.

Finally, I wish to extend my appreciation to my dear family and girlfriend, in Turkey for their deep trust and devotion.

Erhan E YANTERI

Table of Contents

List of Figures

List of Tables

List of Equations

ANALYSIS OF SMALL MUSCLE MOVEMENT EFFECTS ON EEG SIGNALS

I. Introduction

The Brain is the most complex part of the human body and it hasn't been fully discovered and understood yet. Advances in cognitive neuroscience and brain imaging technologies have started to provide us with the ability to uncover the secrets of the human brain and interface directly with it. Electroencephalography (EEG) that is an electrophysiological monitoring method to record electrical activity of the brain was discovered in early 1900s (Teplan, 2002) and it was an important step to discover the secrets of brain activities. Researchers started to understand and formulate how the brain works and reacts to specific events. This significant advancement has encouraged the researchers to achieve one of the biggest dreams of human being which is the communication with machines through thought alone.

Analyzing and formulating of the brain activities provide researchers with the ability to interface directly with the human brain. The needs of people with physical disabilities were one of the biggest motivations to develop Brain Computer Interface (BCI) systems by using this technology. In addition to this, human machine teams have been developed in order to exploit the powerful features of both sides such as the flexibility of the human and the computational power of computers. It is widely accepted in the brain computer interface research community that neurological phenomena are the only source of control in any BCI system (Fatourechi et al., 2007). Because of this reason, brain signals are the key source to develop BCIs and human machine teams and the EEG is a widely used technique to record brain activities.

Developing effective human machine teams requires accurate, fast and reliable communication between human and machine teammates. The Machine side of the team should understand the cognitive state of the human side and act accordingly. Processing of EEG signals makes it possible to recognize the cognitive state. One of the key issues here is to record pure EEG data which contains only the cerebral activity. Since the electrical activity of the brain has a very low amplitude (2-100 µV), the EEG signal is vulnerable to be contaminated by undesired artifacts. In order to make an accurate interpretation of the cognitive state and feed the machine teammate with proper information, it is necessary to deal with these undesired signals.

1.1. Problem Statement

In order to develop effective and autonomous human machine teams, it is vital for the computer agent to understand the cognitive state of the human teammate. Brain activity which is commonly assessed by processing EEG signals is the measure of the cognitive state. When developing teams of humans and computer agents, certain human activities are essential. While interacting with computers, humans perform small motor muscle movements such as operating a keyboard and mouse, manipulating a stick and throttle, or performing touch-screen activities. On the other side, the computer agent needs to know the cognitive state of the human teammate in order to make decisions. EEG signals which are the information source of cognitive state may be affected by small motor movements in fingers, hands and arms. Non-cognitive components in EEG signals are often referred to as artifacts. EEG artifacts may change the characteristics of neurological phenomena and often considered detrimental when trying to determine

operator cognitive state from EEG. Ocular (eye blinks and movements) and muscle artifacts are considered among the most important sources of physiological artifacts (Fatourechi et al., 2007). Large muscle movements, as well as neck, jaw, tongue and shoulder movements are known to generate disruptive artifacts in EEG signals, which reduce the certainty of operator cognitive state. While significant research has explored these large motor effects on EEG and how to remove them (Fatourechi et al., 2007; Vanhatalo et al., 2003; Liu et al., 2016), little is known about small motor effects (hand and finger movements) on EEG. Since the vast majority of human-computer interaction today occurs through keyboard and mouse, knowing/understanding/removing the artifacts generated by hand and finger movement is vital.

The EEG signal is one of the fundamental sources which could be used to determine operator functional state, and knowing operator functional state is a requirement for autonomous decision-making in human machine teams, it is necessary to understand and model the effects of small motor movements on EEG signals. After understanding these affects, we can hope to detect and remove them without eliminating large portions of the EEG signal. Thus, understanding the artefactual effects of small motor movements on EEG signals is important.

This research proposes an experiment to explore small muscle movement effects (artifacts) on the EEG signal. The signal that we are interested in the detrimental effects of the hand/finger movements not the cognitive effects of these movements. Because when we plan or perform muscle movements, it causes changes in the brain activity and these changes in neurological phenomena are not in our area of interest. The primary

expected outcome is evidence about whether, to what extent, and in what ways the EEG signal is affected by small motor muscle movements required for computer operations.

1.2. Research Questions

This research focuses on the observation of the small motor movement effects on EEG signals and their detection. In order to understand these effects scientifically, some questions should be answered. Characterization of hand and finger artifacts, determining their effects on EEG signal and detection of these artifacts form the skeleton of this thesis work. The following questions are asked in order to explore the small muscle movement artifacts and their effects on EEG signals:

Q1. Do small muscle movements (finger and hand movements caused by keyboard and mouse usage) have effects on EEG signals?

In this study, EMG signal (from the right arm) and reference channels which show the exact timing for the keyboard and mouse presses are recorded in addition to the EEG signal. EMG data and reference channels are used to determine the exact times of small muscle movements. The EEG signal can be separated into portions as the data with small muscle artifacts and the data with no artifact by using this additional information. After that separation, it can be possible to analyze these portions and explore the effects of small muscle movements on EEG signal.

Q2. Is it possible to detect small muscle movements (hand and finger movements) from EEG signals, without any reference channel indicating these movements?

In this problem and the thesis work, one of the key points is that we are not simply aiming to detect hand/finger movements. Our aim is to detect these movements by using their artefactual effects, not their cognitive effects.

This question can be considered as a classification problem, since we need to distinguish the data that is affected by small muscle movements and the data that doesn't contain any of these effects. Since we have the reference channels and EMG data, this detection problem can be solved by a supervised machine learning method. The level of small muscle movement effects on EEG signals will determine the success of this machine learning method.

1.3. Assumptions/Limitations

While literature about the EEG artifacts and their characteristics and removal is rich, there is not a comprehensive study about the effects of small muscle artifacts on EEG signal. Due to lack of the former studies about this specific problem, we don't know much about the characteristics of small muscle artifacts and the extent of their effects on EEG signal. This is one of the limitations of this research but the artefactual effects of small muscle movements will be discovered while exploring the first research question.

Some features of EEG signal also cause limitations. EEG signals can easily be contaminated by non-cerebral activities called artifacts. As Klass (1995) states in his study, there are different types of artifacts and they are present in every EEG tracing. It means that it is not possible to record a completely artifact free EEG signal. In our study, the aim is to characterize and detect small muscle artifacts. In order to characterize these effects, we need to compare them with an artifact free EEG recording of the same person.

But it is not possible to record an EEG signal that includes no artifacts in it. Some artifact avoidance methods such us staying in a constant position (in order to avoid muscle artifacts) and not making eye movements have been performed by the participants while recording EEG signal. These artifact avoidance methods reduce the contamination of EEG signal but it is not possible to avoid some of the artifacts such as heart beat, eye blink and other muscle movements. In this study eye movements and eye blinks have been captured by EOG electrodes in order to reject that part of the EEG signal. But for the rest of the EEG recording, it is not possible to get rid of all the artifacts. Because of this reason, we need to assume that the baseline EEG data has no artifact and EEG data with hand/finger movements has only the small muscle artifacts in it.

In general, more positions (EEG channels) mean more information about the cognitive state. In this study EEG signals have been recorded from nine positions on the scalp. Due to equipment capacity, EEG data has only been recorded from these 9 channels and it has been assumed that this is enough to get required information about the cognitive state and small muscle artifacts.

1.4. Contributions

While much research has investigated the relationship between movements of eyes, neck and shoulders and EEG, little is known about hand and finger effects on EEG signal. Large muscle movement artifacts and ocular artifacts have been widely studied and characterized. This study investigates the artefactual effects of small muscle movements and presents a study about a topic that has not been completely

investigated. While ultimately the goal is to detect and eliminate hand/finger artifacts, this work makes contribution only in detection of hand/finger artifacts.

Studies that analyze the effects of finger movements on EEG signal mostly focus on the alpha and beta brain waves to distinguish these effects. We present detailed information about brain waves in the literature review section, showing that alpha and beta waves are strongly related to planning and performing of motor movements. This literature shows that the brain reacts when we plan and perform a motor movement such as hand and finger movements. In these studies, researchers mostly exploit alpha and beta waves in order to detect hand and finger movements and they don't consider their artefactual effects.

On the other hand, this thesis work investigates the artefactual effects of the hand and finger movements on EEG signals (instead of their cognitive effects). According to the results, if the data is heavily contaminated by the hand/finger artifacts (rapid and continuous hand and finger movements with both hands), these artifacts can be detected by analyzing time domain of the EEG signal. If the data contains a small amount of the small muscle artifacts (such as a keyboard press / a simple finger movement), the artefactual effects are not observable. After the time and frequency domain feature extraction and classification, the artifact detection yields good results with high accuracy for highly contaminated data (data heavily includes hand/finger artifacts). On the other hand, the artifact detection accuracy decreased to 64%, for the data that includes only the finger artifacts (index finger keyboard presses). But, after eliminating the blink artifacts the accuracy increased to 72% and this showed that the detection accuracy of the simple

finger artifacts improves, when we eliminated segments of data containing the eye blink artifacts.

While developing human machine teams, computer agent needs to know the cognitive state of the human teammate. It means that computer agent needs the EEG data only represents the neural activity (artifact free EEG data). Because of this, it is important to consider the hand and finger artifacts, if the operator performs intense hand and finger movements.

1.5. Overview

This document is composed of five chapters. Chapter II presents a review of current research focused on EEG artifacts, their detection and removal and analysis of hand and finger movement effects on EEG signal. Chapter III describes the data collection process; processing of EEG data such as feature extraction and time domain to frequency domain conversion and detection of small muscle artifacts. Chapter IV presents the data analysis results, small muscle movement characteristics and performance of their detection. Lastly, Chapter V provides discussion, conclusion and the potential for future work related to this research.

II. Literature Review

This chapter provides a literature review of EEG signals, EEG artifacts and artifact handling techniques. Biological background about brain structure, brain waves and EEG will be given to provide basic knowledge to obtain a better understanding on EEG signals. After the background section, EEG artifacts will be presented in detail. Previous studies about hand and finger movements and their effects on EEG will also be presented in this section.

2.1. Biological Background

The brain is the most complex part of the human body. It is the control center of intelligence, interpreter of the senses, initiator of body movements, and controller of behavior (NINDS, 2012). Scientists and philosophers have tried to discover the secrets and limitations of the brain for centuries, but it is yet to be fully understood. Scientists have learned more about the brain in the past few decades because of the accelerating pace of research in neurological and behavioral science; as well as the development of new research and measurement techniques.

Electroencephalography (EEG) is one of the most important developments in this area. EEG is a kind of imaging technique that measures electrical activity of the brain generated by brain structures. As mentioned by Teplan (2002), the history of the EEG starts in the 1870's. In 1875, Richard Caton presented his findings on electrical phenomena of the exposed cerebral hemispheres of rabbits and monkeys. In 1924, Hans Berger was able to obtain the first human EEG recordings. These improvements have pioneered more and more studies in this field.

The brain is made up of billions of brain cells called neurons, which use electricity to communicate with each other. The combination of millions of neurons sending signals at once produces an enormous amount of electrical activity in the brain. EEG as a monitoring method that records the electrical activity of the brain makes it easier to measure this activity accurately. Quinonez (1998) presents a study about the common applications of the EEG. According to his research, the main diagnostic application of EEG is in the case of epilepsy, as epileptic activity will create clear abnormalities on a standard EEG signal. A secondary clinical use of EEG is in the diagnosis of coma, brain death, tumors and other focal brain disorders.

EEG is also a fundamental and important tool for developing Brain Computer Interface (BCI) systems and evaluation of cognitive status. BCI systems are communication systems that do not depend on the brain's normal output pathways of peripheral nerves and muscles. In these systems, users explicitly manipulate their brain activity instead of using motor movements to produce signals that can be used to control computers, communication devices or physical devices in the real world (Tan & Nijholt, 2010). It is obvious that recording the electrical activity of the brain has opened many new and interesting areas for researchers. As mentioned above, the diagnosis of illnesses, brain computer interface systems and many other areas will continue to attract an increasing number of researchers.

The EEG is one of the most widely used brain sensing methods, since it has several benefits compared to other techniques. The most important benefit of EEG is its excellent time resolution. It means that, it can take hundreds to thousands of snapshots of electrical activity across multiple sensors within a single second. This makes EEG an

10

ideal technology to study the precise time-course of cognitive and emotional processing underlying behavior. This is why EEG signals become the fundamental tool for BCI systems and Human Machine Teams.

In order to understand EEG signals, we need to know the characteristics of the brain signals. Electrical activity of the brain has different patterns that are sinusoidal and these different patterns are called brainwaves. It means that the signal recorded by EEG always includes several brainwaves that are in different frequencies. Secrets of cognitive, affective or attentional states such sleep, attention and wakefulness can be found in these brain waves. EEG signals can represent a wide frequency band (0.5-100 Hz) but the clinical and physiological interest focuses on the frequencies between 0.5 and 30 Hz. The EEG signal is often decomposed into five clinical frequency bands, commonly referred to as waves.

Delta waves (0.5-4 Hz) generally occur while the brain is in very low activity state such as deep sleep (non-REM) and general anesthesia. Theta waves (4-8 Hz) occur in sleep, anesthesia and stress. Alpha waves (8-13 Hz) present while the person is awake, physically and mentally relaxed or in eyes closed position. Beta waves (13-30 Hz) occur while the person is in active thinking, busy or concentration states. Beta waves present strongly while planning or executing motor movements. Gamma waves consist of the frequencies above 30 Hz and the features of these waves are not clear. Because of that, some of the research still doesn't include gamma waves. It is widely considered that gamma waves do not include cognitive processing. This is one of the reasons why the clinical research is mostly focusing on the frequencies up to 30 Hz. Ochoa (2002) also

presents these brain waves in his study as frequency bands of interest. The frequency bands, their range and common association are shown in the Table 2.1.

Table 2.1. EEG Frequency Bands (Ochoa, 2002)

Band	Range	Common Associations
Delta	0.5-4 Hz	Deep sleep; Eye and muscle related artifacts
Theta	4-7 Hz	Emotional Stress; Creative Inspiration; Meditation
Alpha	8-13 Hz	Empty mind; Closed eyes
Beta	13-30 Hz	Active thinking; Attention; Problem solving
Gamma	30 Hz and higher	Blending of multiple brain functions; Muscle related artifacts

2.2. EEG Artifacts

EEG recordings are intended to record the brain activity but these recordings also capture the electrical activities arising from other parts of the body and environment. The non-cerebral components of the EEG signal are termed artifacts.

Fatourechi et al. (2007) define artifacts as undesirable signals that can interfere with neurological phenomena. Since EEG signals are of the order of microvolts (μV), they can easily be contaminated by non-cerebral signals. These interference artifacts can significantly corrupt EEG signal and make its interpretation difficult (O'Regan, 2013). As Klass (1995) presented in his study, artifacts are important for a number of reasons. First, they are present in every EEG recording, and it is not possible to obtain completely artifact-free EEG recording. The artifacts may conceal the actual EEG activity and affect the interpretability of the EEG signal. Moreover, artifacts can lead to false conclusions unless great care is taken to recognize and exclude them from the EEG signal.

An artifact which contaminates EEG signal affects interpretability and may cause false conclusions. Since physiological signals contain valuable information about the

body's physiological and person's cognitive state, artifacts must be handled somehow in order to get maximum benefit from these signals.

Before explaining the artifact handling techniques, it may be beneficial to summarize the sources of artifacts. Klass (1995) classifies the artifacts into three main categories: biological (arising from the subject or patient), technological (arising from the electrode-subject interface, electrode connections or recording equipment) and extrinsic (other equipment connected to the patient; airborne sources, including electromagnetic signals, radio frequency and other environmental phenomena). This categorization is mostly the same in the literature. For example, Files (2011) categorizes the artifact as internal (biological) and external (technological and extrinsic). Similarly, Fatourechi et al. (2007) make the categorization as physiological and non-physiological artifacts. Although they group the artifacts with different names, their main approach in categorization is the same.

The scope of this study is the analysis of small muscle artifacts (hand and finger movements); their characterization, detection and removal. Before focusing on this specific artifact type, it can be helpful to make a brief explanation of artifact types in order to have a wider view. As defined above, artifacts can be divided into two categories as physiological and non-physiological artifacts. Physiological artifacts are mainly ocular (eye movement and blink), muscle (head and shoulder movements, clenching, chewing), cardiac, electro-dermal (sweat artifacts) and glossokinetic (tongue moves) artifacts (O'Regan, 2013). Electrooculography (EOG) (ocular) and electromyography (EMG) (muscle) artifacts are considered among the most important sources of physiological artifacts (Fatourechi et al., 2007a). Non-physiological artifacts are caused by external

effects such as electrode pop or electrode movements and 50/60Hz line artifact. As mentioned above, large muscle movements such as head and shoulder movements are considered in muscle artifacts. Since these movements are close to head and create strong effects, they are even visually visible in EEG signal. On the other hand, small muscle movements don't create strong electrical changes and relatively far away from the head. So, their effects will be less visible in EEG signal. This is the biggest reason that the artefactual effects of small muscle movements have not been examined in the literature.

2.3. Artifact Handling Methods

EEG artifacts are the major problem for interpretation and effective analysis of the EEG signal. In order to develop reliable and effective human machine teams, it is vital to have an artifact-free EEG signal that only represents the cognitive state of human teammate. Because of these issues, dealing with artifacts is critical. There are some artifact handling techniques for EEG signals and these techniques are highly related to each other. Artifact handling techniques can be divided into three parts; artifact avoidance and minimization, artifact rejection and artifact removal.

While obtaining EEG signal from participants, staying in a constant position and avoiding unnecessary movements are appropriate ways of artifact reduction in EEG (O'Regan, 2013). From a data loss and computational perspective, artifact avoidance can be considered as the most ideal technique, since it is assumed that EEG recording contains no artifact if we apply artifact avoidance (Fatourechi et al., 2007a). Unfortunately, it is not possible to avoid all the artifacts. As Klass (1995) states, artifacts present in every EEG signal and some of these artifacts (cardiac artifact etc.) are

unavoidable. Because of that, artifact avoidance and minimization is a good way to reduce the artifact amount, but it is not sufficient to get artifact free EEG signal.

Another artifact handling method is artifact rejection, which is the process of rejecting the trials affected by the artifacts. It means that the EEG signals are evaluated after cutting the time segments of the signal contaminated by the artifacts. This approach is probably the simplest way of dealing with the artifact signals. It has some important advantages over the artifact avoidance approach. First, the experiment for getting EEG signals can take a long time, and the subject does not have to stay completely still. In addition, the secondary cognitive task, resulting from a subject trying to avoid generating a particular artifact, will not present in the EEG signal (O'Regan, 2013).

Researchers have used artifact avoidance and rejection methods to deal with artifacts in earlier studies, but these methods have some drawbacks. These approaches might not acquire sufficient valid data from real experiments, in which eye blinking, swallowing, or other non-neural physiological activities are inevitable (Zhang et al., 2015).

Since these techniques were not able to record the accurate EEG data as best as possible, researchers have focused on another approach to deal with the drawbacks of the earlier methods. Artifact removal that involves the removal of artifact signals has been presented as a solution. A wide variety of techniques have been suggested in the literature; primarily in the areas of epilepsy, evoked and event-related potentials, brain-computer interface and sleep research (O'Regan, 2013). Artifact removal is a much better method than artifact rejection, since this method basically aims to decontaminate the data from the artifacts without rejecting the valuable EEG data, according to O'Regan (2013).

But not all researchers agree with O'Regan's suggestion that artifact removal is better than artifact handling - especially those who are dealing with clinical studies. Because the artifact removal algorithms helps to remove the artifacts but it also changes the data in a way we do not completely understand. Because of that, it is important to determine the area of the study to decide which artifact handling method to use in the study.

The literature on EEG artifact removal is very broad, but researchers still have not agreed on an ideal solution for artifact removal. Types of EEG signals with different characteristics (signal to noise ratio, EEG signal of epilepsy patients etc.), different types of artifacts (muscle, ocular etc.), and lack of not having a common performance measure are three main reasons for this lack of consensus (Urigüen & Garcia-Zapirain, 2015).

Artifact removal can be divided into two types of algorithms: those that perform artifact removal without any additional information and those that perform artifact removal using additional information such as labeled annotations or priori information about the artifact. In addition to these, we should consider the algorithms that combine different methods for artifact removal (hybrid algorithms).

2.3.1. *Artifact Removal Methods*

Linear regression methods were most widely used approaches for artifact removal, especially up to the mid-1990s. As Urigüen and Garcia-Zapirain (2015) define, in linear regression methods, artifacts may be corrected by subtracting a regressed portion of each reference channel from the contaminated EEG. For this to be achieved, we need to know one or more reference channels with the premise that they properly represent all artifact waveforms. It is assumed that each EEG channel is the sum of the artifact-free source signal and a fraction of the artifact that is available through a reference channel.

The simplicity and computationally reduced requirements made this approach popular, especially for EOG artifacts. Since this method requires a reference channel for artifact removal, it can be considered in informed methods.

Filtering is another approach for artifact removal. Simple low-pass, bandpass or high-pass filtering were early classical attempts for artifact removal. However, these methods are not effective when the frequency bands of the EEG and artifact signals overlap (Sweeney et al., 2012). Due to this spectral overlap issue, alternative filtering techniques have been adopted. One of these adopted techniques is adaptive filtering. Sweeney et al. (2012) present a detailed study on this approach. Adaptive filtering assumes that the artifact-free EEG signal and artifact signal are uncorrelated. The filter generates a signal correlated with artifact using a reference channel (that may be obtained by recording EOG, EMG etc.) and the estimate is subtracted from the recorded EEG signal. Adaptive filtering is the most commonly used approach among the filtering method in the artifact removal literature. Since adaptive filtering requires a reference channel to generate signal correlated with artifact, it can be considered as informed algorithm. As stated by Urigüen and Garcia-Zapirain (2015), although these filtering methods need a reference channel (additional information about the artifact), they have the advantage that they can be automated.

Blind source separation (BSS) methods constitute the important portion of the artifact removal algorithms in the literature. BSS is the separation of a set of source signals from a set of mixed signals, without a prior knowledge (or with very little knowledge) about the source signals or the mixing process is a widely used and effective way of artifact removal (Safieddine et al., 2012).

17

Urigüen and Garcia-Zapirain (2015) define the linear mixture of sources model that can be considered the fundamental starting point of BSS approaches. Linear mixture of sources is an approach for artifact removal which adopts the standard assumption that the measured cerebral activity $x(n)$ is the sum of the cerebral activity $s(n)$ and noise $v(n)$.
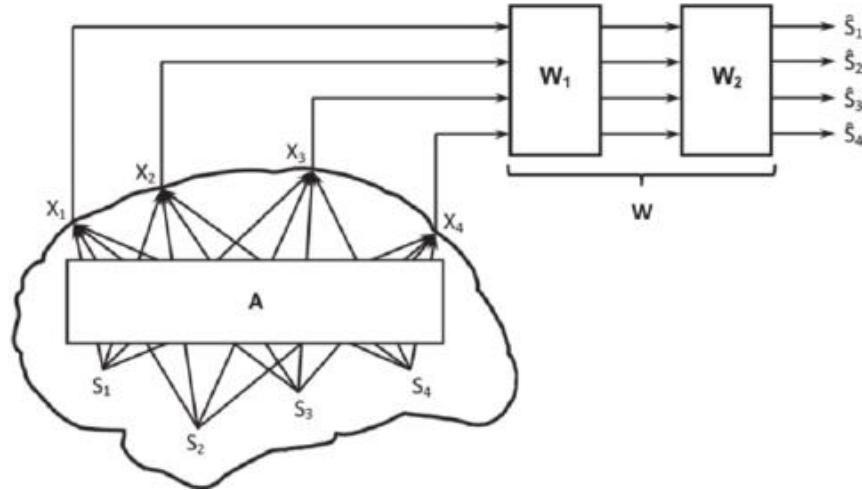
Figure 2.1. Linear Mixture Concept: Combination and Blind Separation of the EEG Sources (Urigüen & Garcia-Zapirain, 2015).

Principle Component Analysis (PCA), Independent Component Analysis (ICA) and Canonical Correlation Analysis (CCA) methods are commonly used BSS methods for artifact removal. ICA is the most widely used BSS artifact removal method.

ICA is a computational method for separating a multivariate signal into additive subcomponents. The starting point for ICA is a very simple assumption that the source components are statistically independent from each other and these components have non-Gaussian distribution (Oja & Hyvärinen, 2000). As emphasized by Oja & Hyvärinen (2000), non-Gaussianity is a key factor for estimating the ICA model. After the ICA

algorithm was introduced, it has largely replaced other methods that are used for artifact removal. The success of the ICA comes from the fact that the brain and artifact signals are sufficiently independent (Urigüen & Garcia-Zapirain, 2015).

BSS methods can be considered as uninformed removal methods, since they don't need a reference channel (additional information about the artifact). Another important common feature about these methods is that they jointly exploit the information provided by all electrodes simultaneously (Safieddine et al., 2012).

Source decomposition methods form another artifact removal approach. As Urigüen and Garcia-Zapirain (2015) defines, in this approach, the problem of finding an artifact-free matrix from the observation matrix is tackled directly by decomposing each individual channel in basic waveforms that can represent either signal or artifact. After decomposition, artifact waveforms are eliminated from each channel individually. Wavelet Transform (WT) and Empirical Mode Decomposition (EMD) are considered under this approach. While BSS methods jointly exploit the information provided by all electrodes simultaneously, EMD and WT process each channel separately (Urigüen & Garcia-Zapirain, 2015).

The techniques mentioned above are the most common artifact removal methods in the literature. BSS techniques are commonly used for artifact removal and ICA is the most widely used method among them. Using the combinations of different artifact removal approaches is another important improvement in this area. Recent studies show that these combinations yield better results than single methods.

When the performances of different artifact removal approaches in the literature are considered, it is concluded that there is no optimal solution (artifact removal

algorithm) for every possible scenario. It is important to evaluate some issues such as the artifact types that are in the data, contamination level, and the type of the EEG signal. The additional information about the artifacts is another criterion for determining the best artifact removal algorithm.

For muscle artifact removal, researchers tend to consider large muscle movements, such as head movements, chewing and clenching. On the other hand, the effect of the small muscle movements, such as hand and finger movements on the EEG signals still remains unexplored.

2.4. Analysis of Hand and Finger Movements

This research focuses on a better understanding of the relationship of hand and finger movement on EEG signals for the purposes of intelligently separating these components from cognitive components in EEG. It is important to understand these effects, since both finger and hand movements are required for human machine interaction.

Lisogurski & Birch (1998) explore classification and differentiation of different sets of muscle movement and show how finger flexions can be identified in continuous EEG signal. Bozorgzadeh et al. (2000) show effects of real and imagined finger movement on EEG and Li et al. (2004) show how EEG signals recorded during finger movement can be distinguished from those during periods of no finger movement. The common point of these studies is that they use alpha and beta frequency bands (8-30 Hz). As Vigneshwari et al. (2013) states in their study, alpha and beta bands includes more information about the cognitive aspects of motor movements. Studies that explore small

muscle motor movement effects use 8-30 Hz (alpha and beta) frequency band. It means that they mostly ignore the artefactual effects of these movements but exploit cognitive effects.

Vigneshwari et al. (2013) analyze finger movements using EEG signal and extract alpha and beta frequency bands by using wavelet transform. In order to discriminate left and right finger movements, they extract different features from alpha and beta bands such as variance and root mean square.

According to literature, it can be concluded that researchers exploit cognitive effects in order to detect finger and hand movements in EEG signal. On the other hand they don't consider the artefactual effects of the hand and finger movements. In this study, we will mostly focus on upper frequency bands to investigate the artefactual effects of these movements and will try to detect these artifacts.

2.5. Summary

This chapter has reviewed EEG, brain waves, EEG artifacts and artifact handling techniques and studies that analyze hand and finger movements using EEG signal. Recent studies show the importance of the EEG for brain computer interfaces and human machine teams. Since it is an inevitable and a fundamental problem for the EEG interpretation, EEG artifacts and their removal methods have been widely investigated.

While developing human machine teams, the EEG signal is a fundamental source of communication for computer agent. In order to accurately interpret the EEG, a computer agent needs to get a pure EEG signal (an EEG signal which only includes the brain activity). On the other hand, human teammates perform simple hand and finger

movements for some operations. These movements can cause the contamination of EEG signal and misinterpretation of human teammate's cognitive state. In this study, we are looking for the artefactual effects of these movements and investigate methods for the detection of them.

III. Methodology

The main objective of this research is to analyze the effects of small muscle movements on EEG signal. Characterization of the hand/finger artifacts and investigating their detectability from the EEG signal form the theme of the analysis process. In this section the following research questions are explored: Q1.) Do small muscle movements (finger and hand movements caused by keyboard and mouse usage) have effects on the EEG signal? Q2.) Is it possible to detect these small muscle movements (hand and finger movements) from EEG signal, without any reference channel indicating hand/finger movement?

These questions aim to determine whether hand and finger movements such as keyboard and mouse activity create artifacts in the EEG signal and develop effective ways of detecting those types of artifacts.

3.1. Domain of Study

In this study, physiological signals have been collected from human subjects. Electroencephalography (EEG), Electromyography (EMG) and Electrooculography (EOG) data have been collected as physiological signals. In addition to the physiological data, digital signals which indicate keyboard and mouse presses have been collected as additional data.

This study has been made as a part of a research ("Small Motor Movement Cognitive Effects") which was approved by Air Force Research Laboratory (AFRL) with human machine subject research protocol: FWR20160127H.

3.1.1. Participants

Five male individuals volunteered to participate in the study. Participants were between 24 and 50 years old (mean age 31.8) and all right handed. The physiological signals have been collected from each subject two times within two different days.

3.1.2. Data Collection

EEG recording system developed by BIOPAC Systems Inc. has been used for the physiological data recording. This system includes the EEG amplifier, EEG electrode cap, EMG and EOG electrodes and Electrode Impedance Checker.

A ribbon cable (100 cm) with connector fans out in the cap to connect to each electrode. The electrode cap's connector arrangement permits the electrode cap to be easily disconnected from the recording amplifiers, allowing the cap to be fitted in one location and used in another. The physiological data recording equipment described above is shown in Figure 3.1.
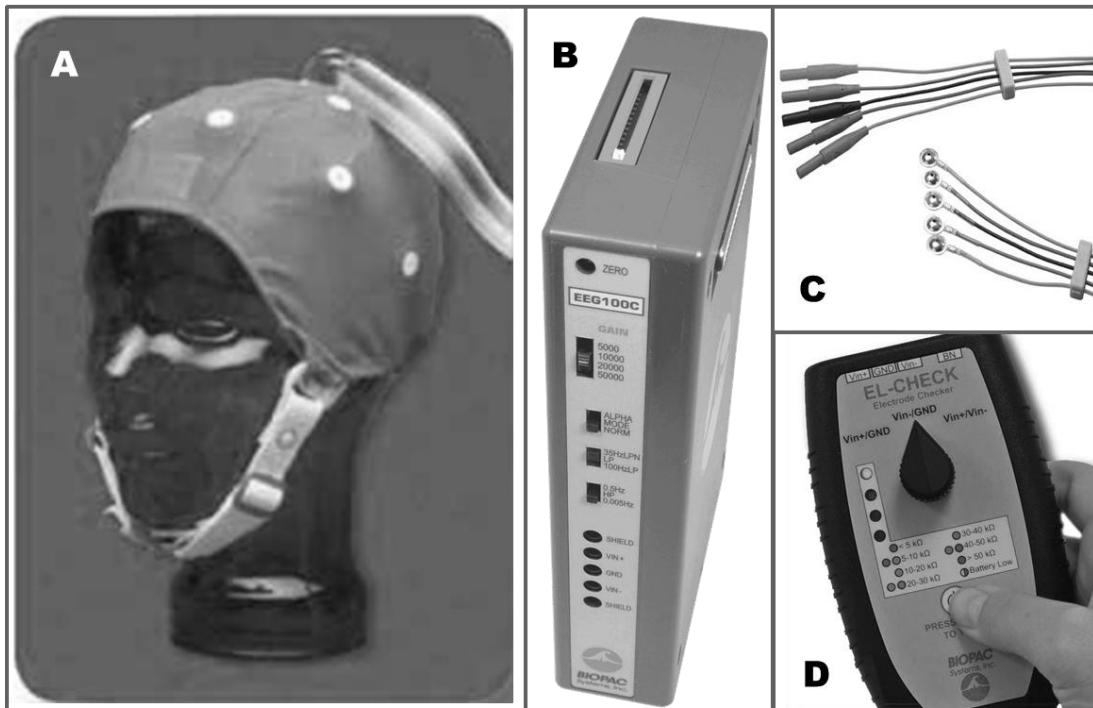
Figure 3.1. Physiological data recording equipment. A. EEG recording cap. B. EEG Amplifier. C. Electrodes that are used for EOG and EMG recordings. D. Electrode Impedance Checker

EEG electrode Cap has the electrodes that are pre-positioned according to the international 10/20 montage that is shown in Figure 3.2.
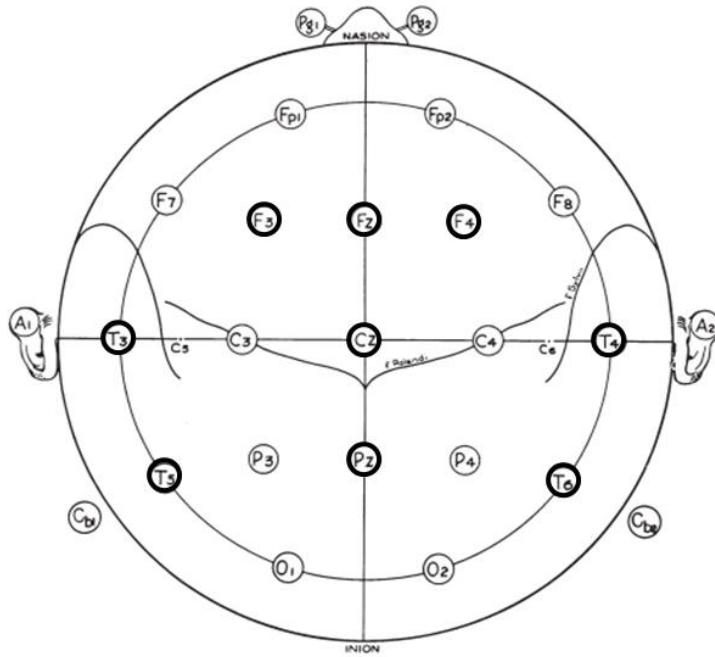
Figure 3.2. The 10-20 International Electrode System (Klem, Lüders, Jasper, & Elger, 1999)

Nine channels have been used for the EEG recording. EEG data has been recorded at a sampling rate of 2000 Hz from the positions of F3, FZ, F4, T3, CZ, T4, T5, PZ and T6 by the EEG electrode cap. Thick circled electrodes in figure 3.2 are the electrodes that we used in this experiment. Two other channels (each channel has two positions: positive and negative) have been used for EOG recordings. One channel has been used to detect vertical eye movements (eye blinks) and the other one was for the

horizontal eye movements. The locations of the EOG electrodes have been shown in figure 3.3.
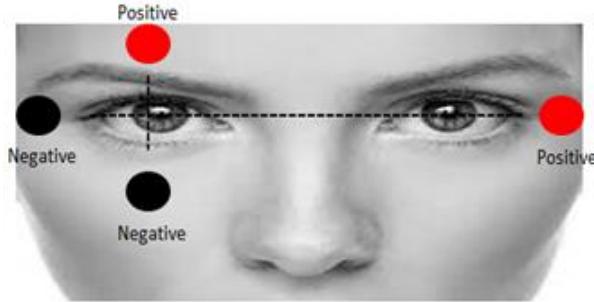
Figure 3.3. EOG electrode locations.

And finally, two channels have been dedicated to EMG recording from the right arm. (It has been used to detect the keyboard presses and mouse movements). The locations of the EMG electrodes (EMG1, EMG2 and the reference electrodes) have been shown in figure 3.4.
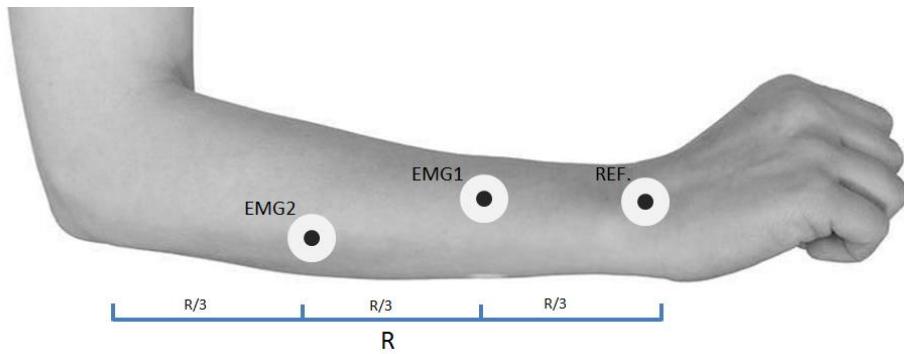
Figure 3.4. EMG electrode locations.

While determining the EMG location, we measured the length from wrist to elbow (R) and this length has been used to determine the EMG electrode locations (1/3 of this length from wrist (R/3) is the location of EMG1 and 2/3 of this length is the location of EMG2).

In the experiment, six different types of conditions have been recorded. Figure 3.4 demonstrates the timeline of the experiment and the duration of each condition.

Figure 3.5. The Timeline of the Experiment.

The first condition (Resting) was the baseline EEG recording. In this condition, the subject was in resting state without causing any artifact and without any cognitive task. The subject has been told to stay in a constant and relaxed position by staring at a constant point on the screen (focusing on the plus sign on the screen) and to clear their minds. Physiological data has been recorded in this state for 60 seconds as the baseline dataset. In the second and the third conditions, the physiological data has been recorded while the subject makes keyboard presses with a single finger. In the second condition, the participant pressed to the left control button with the left index finger. This design was intentionally used to avoid stimulus that would require cognitive processing for the stimulus as well as the keyboard presses. Instead of requesting the participant hit the key after seeing a stimulus, the subject made repeated taps at a participant-determined comfortable, but consistent rate between approximately 1 and 4 taps every 2 seconds (0.5-2Hz) without any stimulus. The aim of this is to reduce the cognitive effects of a visual, auditory, or sensory stimulus on the EEG recording since we are interested in the effects of the small muscle movements on EEG signal. The recording for the second state

27

was made for 45 seconds. The third condition was similar to the second one. This time the subject made repeated keyboard presses with right index finger for 45 seconds. The fourth condition was the imaginary right finger presses. In this state, subject basically did the same thing with the previous condition, but this time he/she made imaginary right finger keyboard presses instead of actually moving their fingers. In the fifth condition, the subject made repeated mouse movement for 60 seconds. These mouse movements were click and drag movements. The participant clicks and drags the mouse from the left to the right and returns the mouse to its original position again. He/she performs this movement several times (in a participant-determined comfortable, but consistent rate between approximately 1 and 4 mouse movements every 2 seconds) for 60 seconds. In the final state, the subject used keyboard and mouse at the same time with a stimulus that shows how he/she is performing. The subject used two fingers on the left hand to alternately press the A and D keys and at the same time he/she made mouse movements with the right hand. This data was considered as fully contaminated by the small muscle movement effects and compared with the baseline data.

3.2. Preprocessing of the EEG Data

In order to perform feature extraction and artifact detection from EEG signal, it is essential to preprocess the raw EEG data to improve the performance of the analysis. In this thesis work, an open source toolbox called EEGLAB provided by SCCN lab, running under the cross-platform MATLAB environment (The Mathworks, Inc.) has been used for both preprocessing and some parts of the EEG analysis.

The first step of preprocessing was the conversion of the data. The EEG signal was recorded using the integration of "AcqKnowledge Data Acquisition and Analysis Software" and "PsychoPy" software. The recorded data (9 channels EEG, 2 Channels EOG, 2 channels EMG and digital reference channels) has been converted to the appropriate format (.mat) for MATLAB and it has been formatted in order to make the analysis easier. This raw EEG data (.mat file) was including all the recording. Because of that, we divided this data into meaningful pieces by using MATLAB programming and saved each experiment condition data separately (For example: Participant_1 Day_1 Resting State).

The data analysis functions available in EEGLAB which includes data filtering, data epoch extraction and data resampling were used in this thesis work for the preprocessing of the collected EEG data. Since the EEG data has been acquired with a high sampling rate (2000 Hz), in some parts of the analysis the data has been down sampled. Down sampling reduces the file size and speeds up the subsequent processing steps and this was necessary for some computationally intense analyses such as time-frequency power maps. In order to facilitate investigating task-related changes in the EEG, we cut the continuous data into segments surrounding particular events (Finger and hand movements). These epochs includes the signal with hand/finger movements and the signals with no small muscle movements. The usage of the epochs is described in Section 3.3 in detail.

For Preprocessing, we first imported the data into EEGlab and made channel localization. Channel Localization is important for ICA and investigating each channel separately. Figure 3.6 show how we made the channel localization in EEGlab.

The EEG data we recorded includes 60 Hz line noise and we needed to get rid of this first. Figure 3.7 demonstrates the 60 Hz line-noise-effect in frequency-power spectrum.
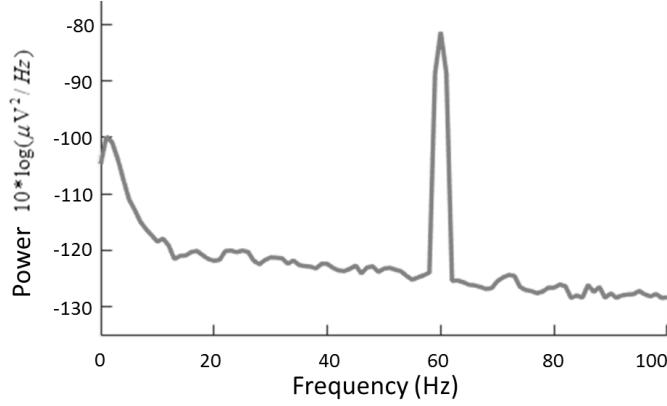
Figure 3.6. Frequency-Power Spectrum. (Resting State Recording of Participant One).

In order to remove this artifact, we examined three different approaches. These approaches were notch filtering, filtering with Cleanline toolbox (It is a sinusoidal artifact removing toolbox in EEGlab) and ICA (Removing components that includes 60 Hz artifact). The results of these approaches will be presented in the analysis section (Section 4.1).

These preprocessing steps are important to get better and more accurate results. After obtaining the proper signal for analysis, the next step was the feature extraction.

3.3. Physiological Feature Extraction

Acquisition of large amount of data is obtained by measuring electrical activity of the brain through EEG electrodes. In this study, 9 electrodes have been used for recording EEG signals with a sampling frequency of 2000 Hz. In order to detect the artefactual effects of hand/finger movements in EEG signal, it is essential to find features that can be

30

helpful to distinguish these effects. "Features" are the values which define some relevant properties of the acquired signals and combined as "feature vector". Hence, feature extraction is an operation which converts one or several signals into a feature vector.

Determining and obtaining required features from EEG signals is an important step. Since we are trying to characterize the finger and hand movement effects on the EEG signals, we need to extract the features that represents these effects. Many extraction techniques have been proposed and studied in the literature to represent EEG signals, such as wavelet transform, power spectra and adaptive autoregressive (Vigneshwari et al. 2013). In this study, we extracted the features that may obtain valuable information about the effects of the finger and hand movements.

3.3.1. Wavelet Transform

Wavelet transform was one of the feature extraction methods we used in this thesis work. Wavelet transform is a time-resolved frequency decomposition of EEG data. It is a useful decomposition technique for our case, since the frequency-domain representations of EEG data such as Fourier transform have some limitations. They are unable to visualize the changes in frequency structure over time. On the other hand, a wavelet transform can provide us with the frequency of the signals and the time associated to those frequencies and it is a good way to visualize and decompose EEG signals into measurable component events.

We decomposed the EEG data into the sub frequency bands by using the discrete wavelet transform.
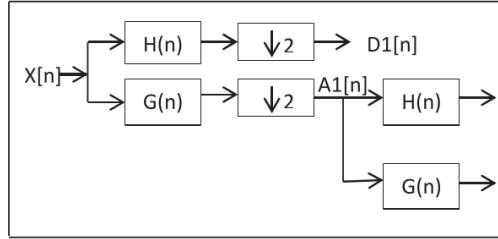
Figure 3.7. Wavelet decomposition tree

Figure 3.7 illustrates the discrete wavelet transform tree. The DWT is computed by applying successive low-pass and high-pass filters to the discrete time-domain EEG signal and this figure is called the Mallat tree decomposition (Polikar, 1994). In Figure 3.7, the EEG signal is denoted by the sequence $X[n]$, where n is an integer (the voltage value from one channel at a specific time). The low pass filter is denoted by $G(n)$ and the high pass filter is denoted by $H(n)$. At each level, the high pass filter produces detail information, $D[n]$, while the low pass filter associated with scaling function produces coarse approximations, $A[n]$.

Wavelet transform was used to create time-frequency power plots in order to demonstrate the changes in different frequencies by the effects of hand/finger movements. It was also used to convert time domain EEG signal into the frequency domain. The output of the wavelet transform was used to extract features from frequency domain data.

3.3.2. *Extracting Different Frequency Bands by Using FIR Filter*

In the literature, studies which focus on designing BCI systems based on EEG use alpha (8-13 Hz) and beta (13-25 Hz) waves (frequency bands) as information sources of the systems. As Gunaydin and Ozkan (2010) state, alpha and beta waves contain more information about small muscle motor movements (such as hand and finger movements).

32

This information represents the cognitive aspects of the small muscle motor movements. It means that, these studies exploit the cognitive changes in the brain in order to distinguish muscle movements (such as hand and finger movements). In this study our aim is not to analysis the cognitive effects of the small muscle movements, but the artefactual effects of the muscle movements themselves on the EEG signal. Because of that, instead of using these waves (frequency bands), we need to use other frequency bands that don't contain cognitive features (or contain very little cognitive information).

In this study, we used FIR filters in order to extract different frequency bands. EEGlab toolbox has been used for this purpose. EEGlab has some filtering tools and Basic FIR filter tool (pop_eegfiltnew function) has been used for band pass filtering. We extracted 3 different frequency bands; 10-25 Hz, 30-50 Hz and 50-100 Hz. We extracted 10-25 Hz frequency band since it includes alpha (8-13 Hz) and beta (13-25 Hz) waves. 30-50 Hz and 50-100 frequency bands were extracted to investigate the artefactual effects of small muscle movements.

3.3.3. *Feature Extraction from Time Domain Data*

We acquired 9 channel EEG signal at 2000 Hz sampling rate (with additional channels) and this means a large amount of data. Feature extraction is a kind of dimensionality reduction and it is important while dealing with large amount of data in pattern recognition. Features are values that represent some properties of the acquired signal.

In this study, time and frequency domain features have been extracted from each channel in order to investigate artefactual characteristics of the finger and hand movements. Recorded EEG signal was preprocessed and different frequency bands were

extracted as we mentioned earlier. After that process, we divided signal into 0.5-second-chunks (1000 samples per chunk) and important time domain features have been extracted from each of these chunks. Figure 3.9 shows the flow chart of the feature extraction from time domain data.
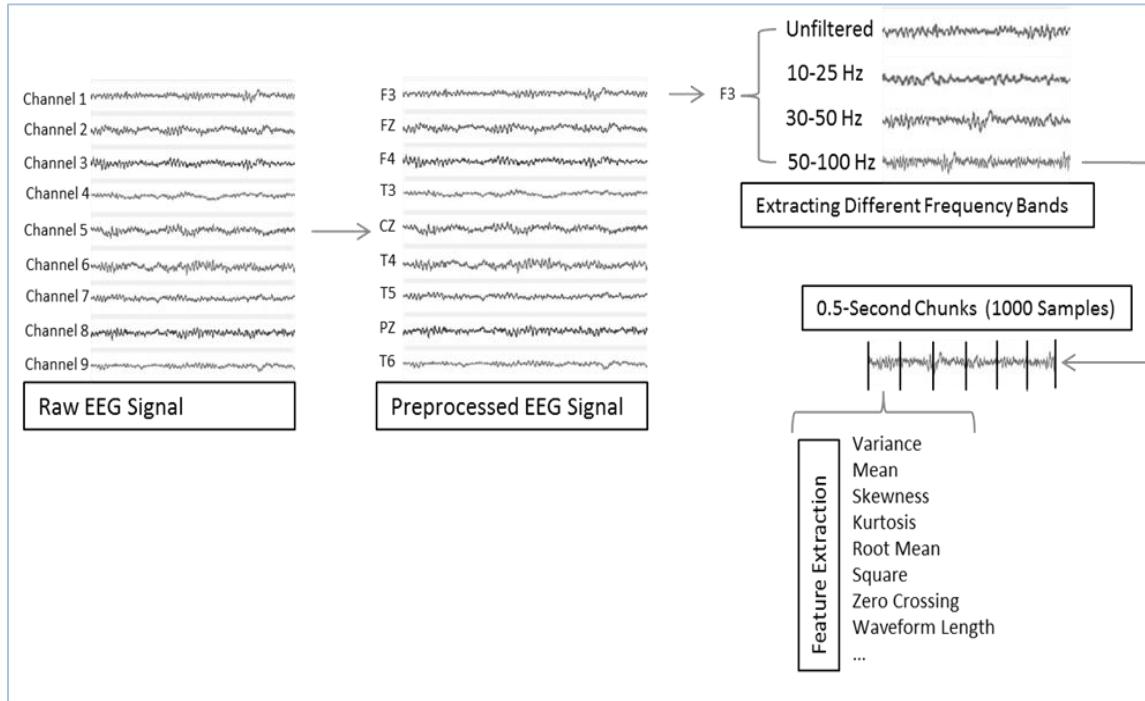
Figure 3.8. Flow Chart of Preprocessing, Frequency Decomposition and Feature Extraction of 9 Channel EEG Data.

There are different features that can be used for feature extraction. The features that were investigated in this thesis work have been listed and explained below.

- o Integrated EEG
- o Root Mean Square
- o Mean Absolute Value
- o Variance

- o Waveform Length

- o Zero Crossing

- o Mean

- o Skewness

- o Kurtosis

3.3.3.1. *Integrated EEG*

Integrated EEG (IEEG) is calculated as the summation of the absolute values of the EEG signal amplitude. It can be expressed as shown in Equation 1.

$$IEEG = \sum_{n=1}^{N} |X_n| \tag{1}$$

In this equation, N represents the number of the samples of the specific channel; X_n represents the voltage value of EEG data at a specific time.

3.3.3.2. *Root Mean Square*

RMS is known as the quadratic mean. In statistics, the root mean square (RMS), also known as the quadratic mean, is defined as the square root of the arithmetic mean of the squares of a set of numbers. RMS is a useful feature when there are positive and negative variations, such as EEG signal. The formulation is expressed as shown in Equation 2.

$$RMS = \sqrt{\frac{1}{N} \sum_{n=1}^{N} X_n^2} \qquad (2)$$

In this equation, N represents the number of the samples of the specific channel; X_n represents the voltage value of EEG data at a specific time.

3.3.3.3. *Mean Absolute Value:*

Mean Absolute Value can be calculated by taking the average of the absolute value of EEG signal. As Vigneshwari et al. (2013) supports, it is an easy way for detection of muscle contraction levels. It is defined as shown in Equation 3.

$$MAV = \frac{1}{N} \sum_{n=1}^{N} |X_n| \qquad (3)$$

3.3.3.4. *Variance*

Variance of EEG (VAR) uses the power of the EEG signal as a feature. Generally, the variance is the mean value of the square of the deviation of that variable. It can be expressed as shown in Equation 4.

$$VAR = \frac{1}{N-1} \sum_{n=1}^{N} X_n^2 \qquad (4)$$

3.3.3.5. *Waveform Length*

Waveform length (WL) is the cumulative length of the waveform over the time segment. WL is related to the waveform amplitude, frequency and time. It is defined as shown in Equation 5.

$$WL = \sum_{n=1}^{N} | X_{n+1} - X_n |$$
(5)

3.3.3.6. *Zero Crossing*

Zero crossing (ZC) is a point where the sign of a mathematical function changes (e.g. from positive to negative), represented by a crossing of the axis (zero value) in the graph of the function. That means that it represents the number of times that the amplitude value of EEG signal crosses the zero y-axis. This feature provides an approximate estimation of frequency domain properties. It is formulated as shown in Equation 6.

$$ZC = \sum_{n=1}^{N-1} | \operatorname{sgn}(X_n) - \operatorname{sgn}(X_{n-1}) |$$
(6)

Where $sgn(X_n) = 1$ when $X_n > 0$ and $sgn(X_n) = 0$ in other conditions.

3.3.3.7. *Mean*

Mean is calculated as the usual average of the EEG signal amplitudes. It can be expressed as shown in Equation 7.

$$Mean = \frac{\sum_{n=1}^{N} X_n}{N} \tag{7}$$

3.3.3.8. *Skewness*

Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the center.

3.3.3.9. *Kurtosis*

Kurtosis is a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution. That is, data sets with high kurtosis tend to have heavy tails, or outliers. Data sets with low kurtosis tend to have light tails, or lack of outliers.

3.3.4. Feature Extraction from Frequency Domain

Feature extraction from the frequency domain EEG data is important to observe and investigate another dimension of the data. In this study, time domain EEG data was converted into the frequency domain data by using wavelet convolution and Hilbert transform. Frequencies form 70 Hz to 100 Hz were convolved with proper wavelets and their power values was obtained for each channel.

As a result, we had the frequency power values of 31 different frequencies (70-100 Hz) for each time points and each channels. After that, this frequency domain data was divided into 0.5 second chunks and important features were extracted from each chunk. One of these features was the average power. Average power is calculated by taking the mean of all frequency power values in that chunk. The other feature was the maximum power. Maximum power was calculated by taking the maximum power value of averaged frequency powers. The frequency that has the maximum frequency power

was another feature. Minimum frequency power and Root mean square were the other features that we extracted from the frequency domain. Figure 3.9 demonstrates the feature extraction process from the frequency domain (Extraction of the Maximum Frequency Power).
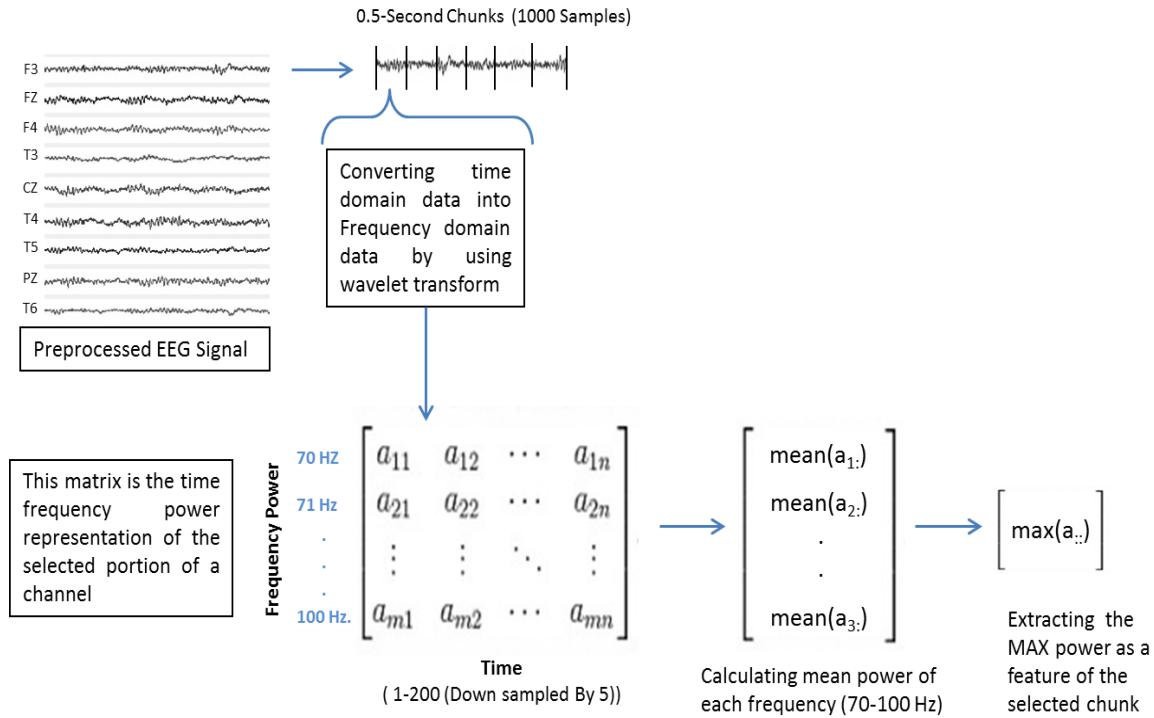
Figure 3.9. Flow Chart of Feature Extraction from the Frequency Domain.

3.4. Analysis of EEG signals

In order to investigate the first research question, we used some analysis methods to observe how EEG signals changes during the hand and finger movements. In order to visualize these changes, we compared the event related signals (signals that include hand and finger movements) to the baseline signal. (The baseline signal refers to EEG signal

that doesn't include any hand or finger movements). As we mentioned, we collected EEG data from subject for six different conditions; resting, left finger keyboard press, right finger keyboard press, right finger imaginary press, mouse movement and lastly keyboard press + mouse movement with stimulus (we named the last state as the activity state).

First we compared the first and the last conditions. We assumed that one of them (resting state) includes no finger and hand movement effects and the other one (activity state) completely includes these effects. Since the activity state heavily contaminated by hand and finger artifacts, it was easier to observe the effects of these movements. After that, we analyzed other conditions such as left and right finger press and right finger imaginary press conditions. We plotted the frequency power density plots, in order to investigate the effects in the frequency domain. After that by extracting features from the time and frequency domain data, we aimed to find useful features to differentiate hand and finger movement artifacts.

In order to compare the first and the last conditions, we converted the 60 second time series EEG data of the first condition (resting) and 60 second data of sixth condition (keyboard press and mouse movement activity) into the frequency domain by power spectral density plots. This is a kind of visual analysis in order to investigate which frequency bands differentiate between two conditions. Figure 3.10 demonstrates the process. In resting state there are no hand/finger artifacts, but the participant creates hand/finger artifacts by both hands in the activity state.
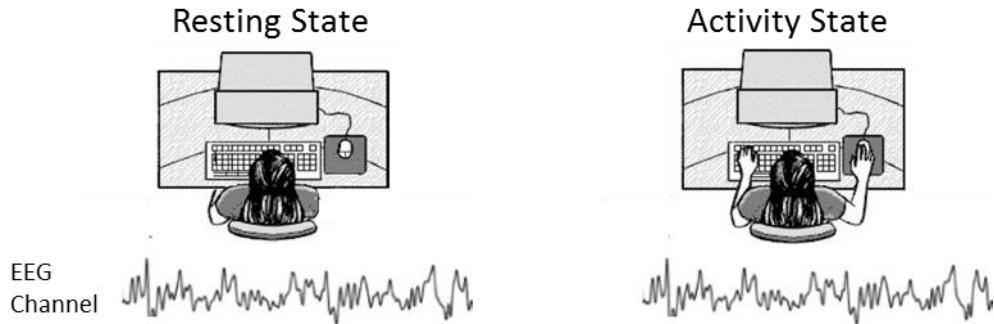
Figure 3.10. The Resting and the Activity States

We needed to use another approach for the other conditions. We had two separate data for the resting and activity conditions, but it is not the same for other conditions. Because of that, we needed to separate data into segments. One group of these pieces doesn't include finger and hand artifacts and the other group includes these artifacts. Figure 3.11 shows how this process works.
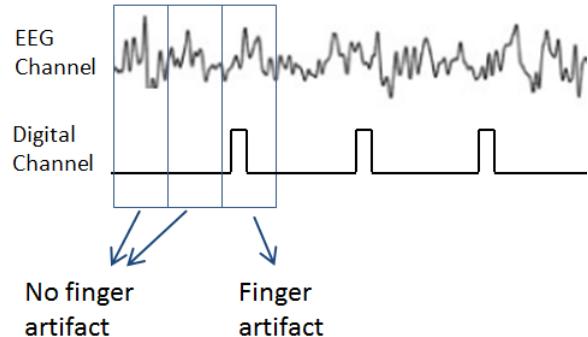
Figure 3.11. Separating Artefactual and Non-Artefactual Portions of EEG Signal.

We first divided the data into the segments (0.5-second chunks), after that we used digital channels, in order to determine the chunks have the finger artifacts or not. Any chunk which contains a digital indication of a finger movement is considered to be a

finger movement chunk. As a result we grouped these artifacts as chunks with finger artifact and chunks with no finger artifact.

Finding the eye blink artifact was another important step in this study. In order to find eye blink artifacts, we used VEOG channel that shows the vertical eye movements (blinks). The VEOG data was low pass filtered (20 Hz) to remove jagged edges and the rest of the channel has been investigated to find amplitudes bigger than blink voltage (it is a values that we determined visually. The value of it is 50 μV). As a result we were able to find when the participant blinked.

3.4.2. *Detection of the Hand and Finger Movements Effects*

After characterizing the effects of the hand and finger movements, the second important issue is to detect these effects without any reference channel. While detecting hand and finger artifacts, we didn't use neural changes, but the artefactual effects of these movements. We can consider this detection process as a classification problem, since we are trying to classify the data that is affected by the small muscle movements and the data that doesn't include any of these effects.
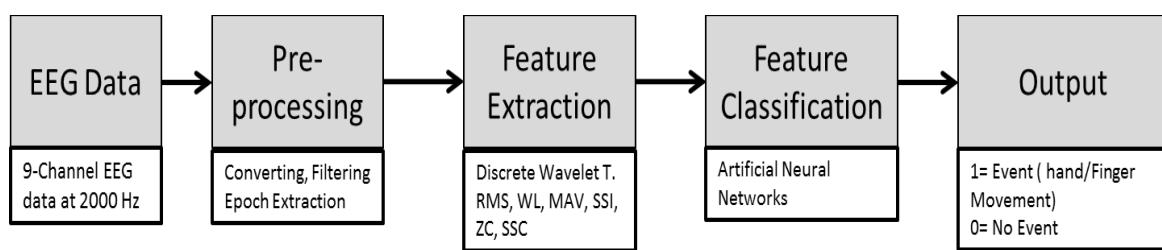

Figure 3.12. Block Diagram Of the Classification

Figure 3.12 illustrates the process of classification. As a result of this process, we can detect the data that is affected by the hand and finger movements. In the literature, various classification algorithms were used for the classification of the EEG signal. These

are linear classifiers (Linear Discriminant Analysis- LDA, Support Vector Machine-SVM), non-linear Bayesian classifiers, Linear Discriminant Analysis and Neural Networks. As Varghese (2009) states, the main drawback of LDA is its linearity that can provide poor results on complex non-linear EEG data. Because of that we used 3-layer (input, hidden and output layers) feed-forward neural network models for classification. Figure 3.13 demonstrates a sample for a 3-layer feedforward neural network.
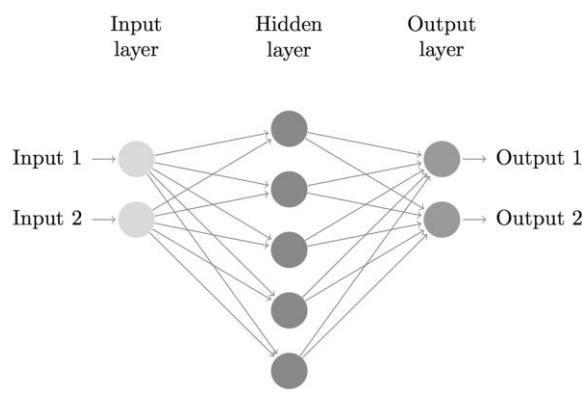
Figure 3.13. Basic 3-Layer Feed Forward Neural Network

Since the feature numbers and feature matrix sizes were different for different classification conditions (resting-activity classification, right/left index finger press classification), different numbers of input, hidden and output nodes have been used. But the basic structure of the networks was the same. The detailed structure of different classifications will be presented in the Analysis part. The Neural Network Pattern Recognition tool of MATLAB (`nnstart`) was used to create these structures and train/test the neural networks. These neural networks were fully connected feed-forward networks with sigmoid hidden and softmax output activation functions. The data was divided into two sets as training (70%) and test (30%) data. The 70% of training data was

43

used to feed the network and 30% of it was used for validation. After the training session, test data was tested by using the trained model. We will explain how we divide the data into the training and test sets in the Analysis section.

3.5. Evaluation

In order to evaluate the success of the detection and the classification of small muscle movement artifacts, we need some performance measures. The results of the classification are evaluated and presented by the confusion matrixes. These confusion matrixes (error matrixes) allow visualization of the performance of the classification model. Each column of the matrix represents the instances in a predicted class while each row represents the instances in an actual class (or vice versa). The figure 3.14 represents a basic structure of a confusion matrix.

		Predicted Condition	
	Total Population	Predicted Condition Positive	Predicted Condition Negative
True Condition	Condition Positive	True Positive	False Negative (Type II Error)
	Condition Negative	False Positive (Type I Error)	True Negative

Figure 3.14. Basic Structure of a Confusion Matrix

Accuracy is another representation of performance in the classification problems. It shows how often the classifier is correct. The accuracy can be calculated by using the confusion matrix. The calculation of the accuracy is shown in the Equation 8.

$$Accuracy = \frac{TruePositive + TrueNegative}{Total\_Population} \tag{8}$$

The Receiver Operating Characteristic (ROC) Curve is another evaluation tool for our classification. ROC is a plot of values of the False Positive Rate (FPR) versus the True Positive Rate (TPR) for all possible cutoff values from 0 to 1. Area under ROC curve indicates the performance of the classification.

3.6. Summary

This chapter described the analyzing of the hand/finger artifacts and characterization of them, detection of the hand/finger artifacts by using time and frequency domain EEG data. Characterization of these artifacts was made by the frequency power density plots and feature extractions. After the characterization of these artifacts, we used proper frequency bands and features for the detection. The detection of these artifacts was made by using neural networks. Analysis results and the detection performances will be reported in the following chapter.

IV. Analysis and Results

In this chapter, the effects of small muscle motor movements are investigated by analyzing recorded EEG data in time and frequency domain. Features which are extracted from time and frequency domain are used to find evidence of small muscle movement effects. The features that include valuable information about the hand and finger artifacts are used to train the neural network models and detection of these artifacts has been made by using the trained models. Unlike Li et al. (2004) and Vigneshwari et al. (2013)'s studies, this thesis work investigates the artefactual effects of the hand and finger movement. Because of that, features are extracted from upper frequency bands instead of exploiting alpha and beta bands.

These analyses and classification models are made to find answers to our research questions: Q1.) Do small muscle movements (finger and hand movements caused by keyboard and mouse usage) have effects on EEG signals? Q2.) Is it possible to detect small muscle movements (hand and finger movements) from EEG signals, without any reference channel indicating hand/finger movement?

4.1. Removing 60 Hz Line Noise

EEG signal that we recorded includes 60 Hz line noise. This noise can be seen even in the raw EEG signal. Figure 4.1 shows the raw EEG signal with 60 Hz line noise and the frequency power spectrum on one channel of that EEG signal.
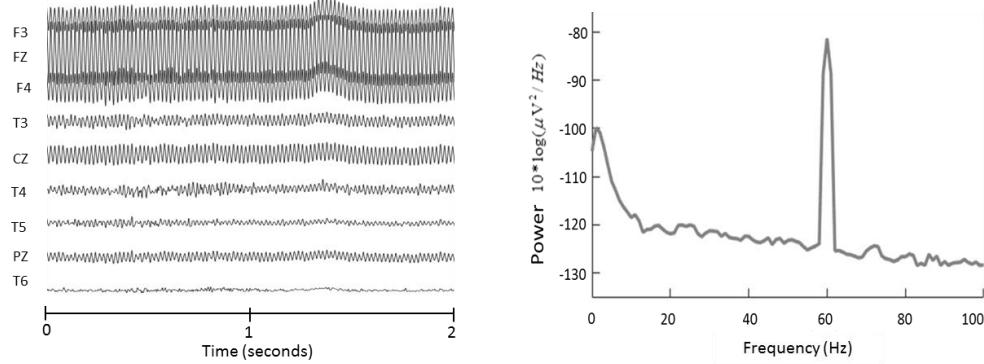
Figure 4.1. Raw EEG Signal with 60 Hz Line Noise and Frequency Power Plot of

Channel CZ.

In order to remove this noise, 3 different approaches were investigated. The first one was applying notch filter between 58-62 Hz. The second option was filtering the signal with Cleanline toolbox (It is a sinusoidal artifact removing toolbox in EEGlab). And the last option was applying ICA and removing components that include 60 Hz line noise. Notch filtering creates band holes and distorts frequencies around the notch frequency, but according to the results, it is the best option among these three approaches. Figure 4.2 shows the results of 3 different approaches by frequency power plots.
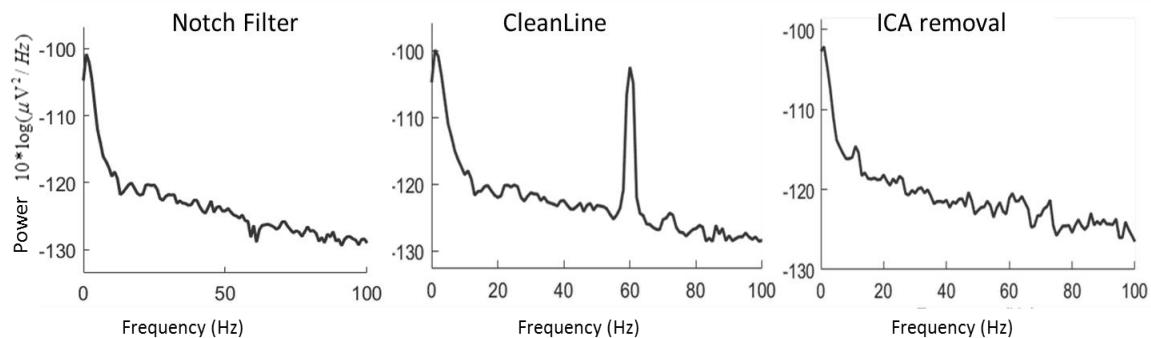

Figure 4.2. Frequency Power Spectra of EEG Signal after Notch Filtering (58-62 Hz),

CleanLine Filtering and ICA Removal

47

CleanLine reduces the effect of 60 Hz Noise, but the power of 60 Hz noise still remains in the signal. ICA is used to decompose the signal into independent components. For 60 Hz line noise removal, we removed the components that include 60 Hz signal and rest of the components was composed again. From figure 4.2, it can be seen that the 60 Hz line noise is removed by ICA, but ICA also affected the other frequencies because of the removed components. In addition to this, ICA is designed to remove artifacts, and if we are not careful, we might end up removing the artifacts we hope to investigate. As a result, we decided to use the notch filtering to remove the 60 Hz. Line artifacts.

4.2. Frequency Power Spectra

4.2.1. Comparing Resting and Activity Conditions

In first part of this thesis work, we compared the resting and the activity conditions. The resting state contains no hand and finger movements and the activity state contains both finger and hand movements (The left hand for keyboard presses and the right hand for the mouse movements). We started comparison by plotting the frequency power spectra. Figure 4.3 shows the frequency power spectra of the resting and the activity conditions of participant one. Each plot includes the power spectrum of both activity and resting conditions.
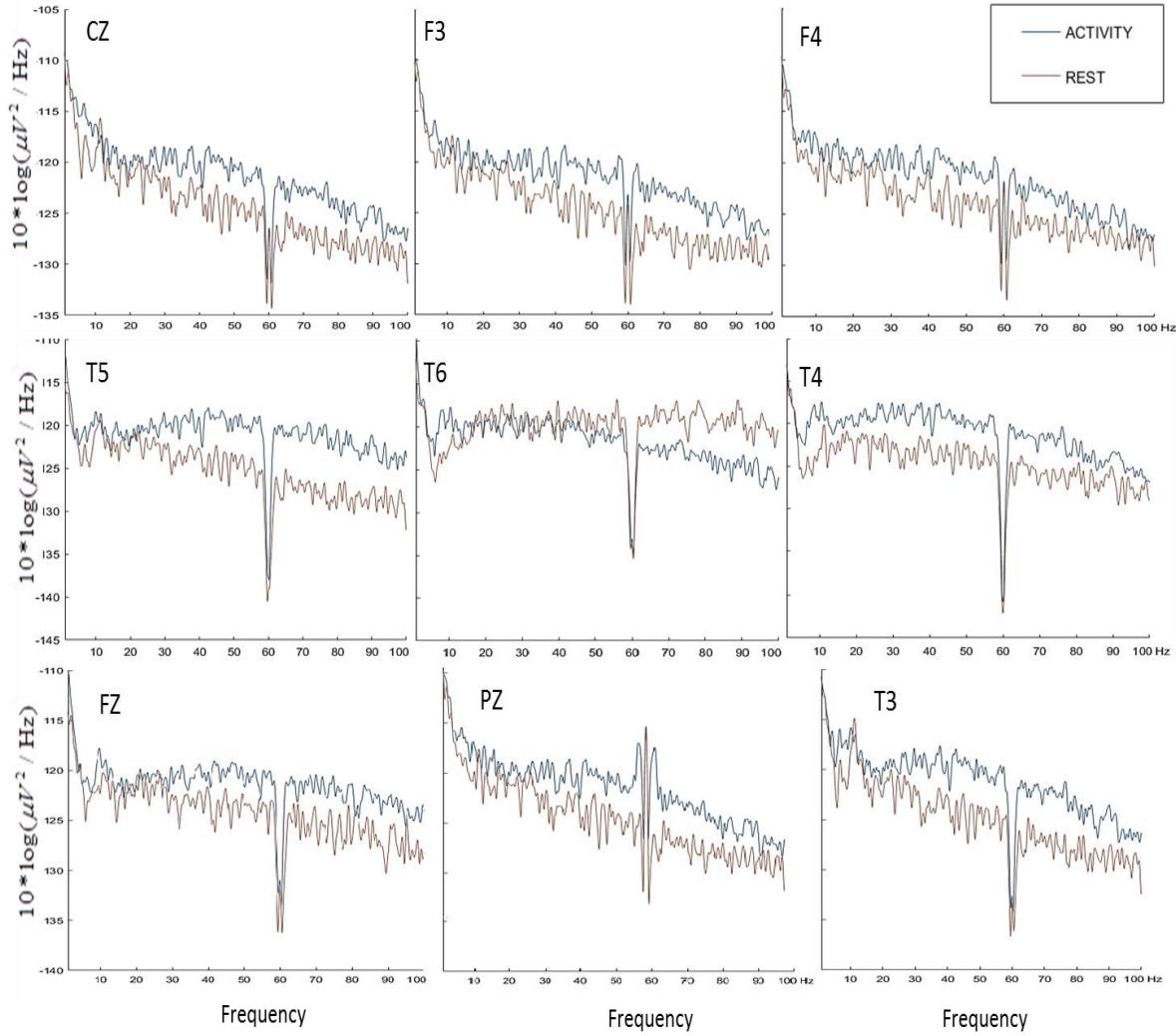
Figure 4.3. Frequency Power Spectra of Resting and Activity Conditions (Data has been Taken from the Participant-One Day-One Recording).

According to frequency power spectra of each channel, it can be observed that the power of upper frequencies is relatively bigger in the activity state. This can't be the evidence for the effects of hand and finger movement, but it demonstrates that performing activity instead of staying in constant position affects the power spectrum of EEG signal. We know that lower frequency bands (<25 Hz) include more cognitive components. On the other hand, upper frequencies (>25 Hz, gamma wave) contain less

cognitive components. As Whitham et al., (2007) suggest, EEG recording above 20 Hz could be in many cases an artifact of electromyography activity. Power spectra in figure 4.3 show that the frequency power of the activity state is bigger than the resting state, especially above 25 Hz. This can be an indicator of the artefactual effect of hand and finger movements. This situation is the same for all 5 participants (Activity state has relatively bigger power for upper frequencies).

4.2.2. *Right Finger Press and Right Finger Imaginary Press Conditions*

We recorded two different conditions for right finger keyboard press in order to explore the effects of right index finger movement on the EEG signal. One condition contains right finger movements and the other contains imaginary right finger movements. We plotted frequency power spectra of both conditions in order to investigate finger movement effects. Although the frequency power spectra of resting and activity states generated visual difference between two conditions, the power spectra didn't generate any visual difference between the right finger movement and the right finger imaginary movement states. Figure 4.4 demonstrates the power spectra of these two conditions (45 seconds recording of each conditions) for 3 channels.
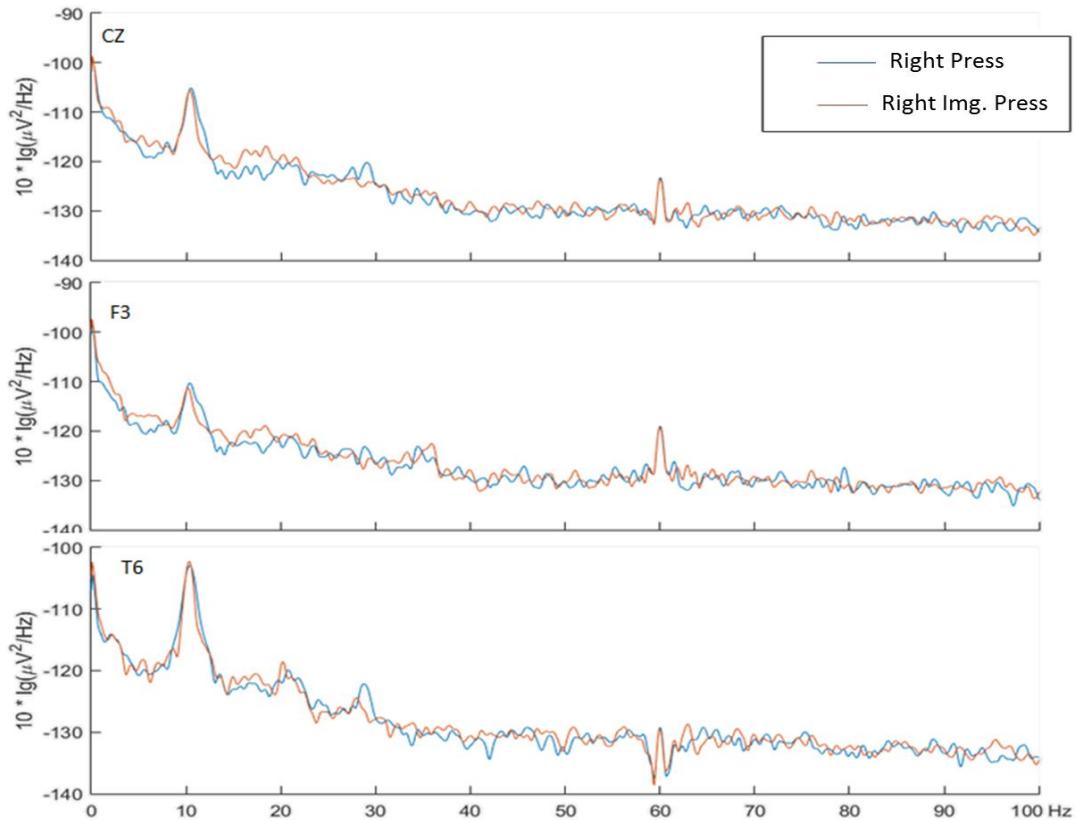
Figure 4.4. Frequency Power Spectra of Right Finger Press and Right Finger Imaginary Press Conditions (Data has been Taken from the Participant-One Day-One Recording).

The power spectra show that the finger movements don't create any observable effects in frequency spectrum as in the previous comparison. We need to try other options in order to find an evidence for the artefactual effects of finger movements on the EEG signal.

4.3. Feature Extraction and Creating Feature Matrixes

4.3.1. Resting and Activity Conditions

We observed that the EEG data which is recorded while performing hand and finger movements has more power in upper frequencies than the EEG data which is recorded

without making any muscle movements. In this part of the thesis work, we extracted features from time and frequency domain EEG data.

The data from resting and activity conditions has been band-pass filtered into different frequency bands (10-25 Hz, 30-50 Hz and 50-100 Hz) by using FIR filtering. After filtering EEG data, each data portion (activity and resting conditions) has been divided into 0.5-second chunks and time-domain features have been extracted from each chunk as we described in the methodology part (Figure 3.8). As a result, we obtain two groups of features. One of the groups was the features extracted from resting condition and the other one was the features of the activity state. The same features from two groups were visually compared by boxplots. Figure 4.5 shows boxplots of 8 different features that belong to two different groups.
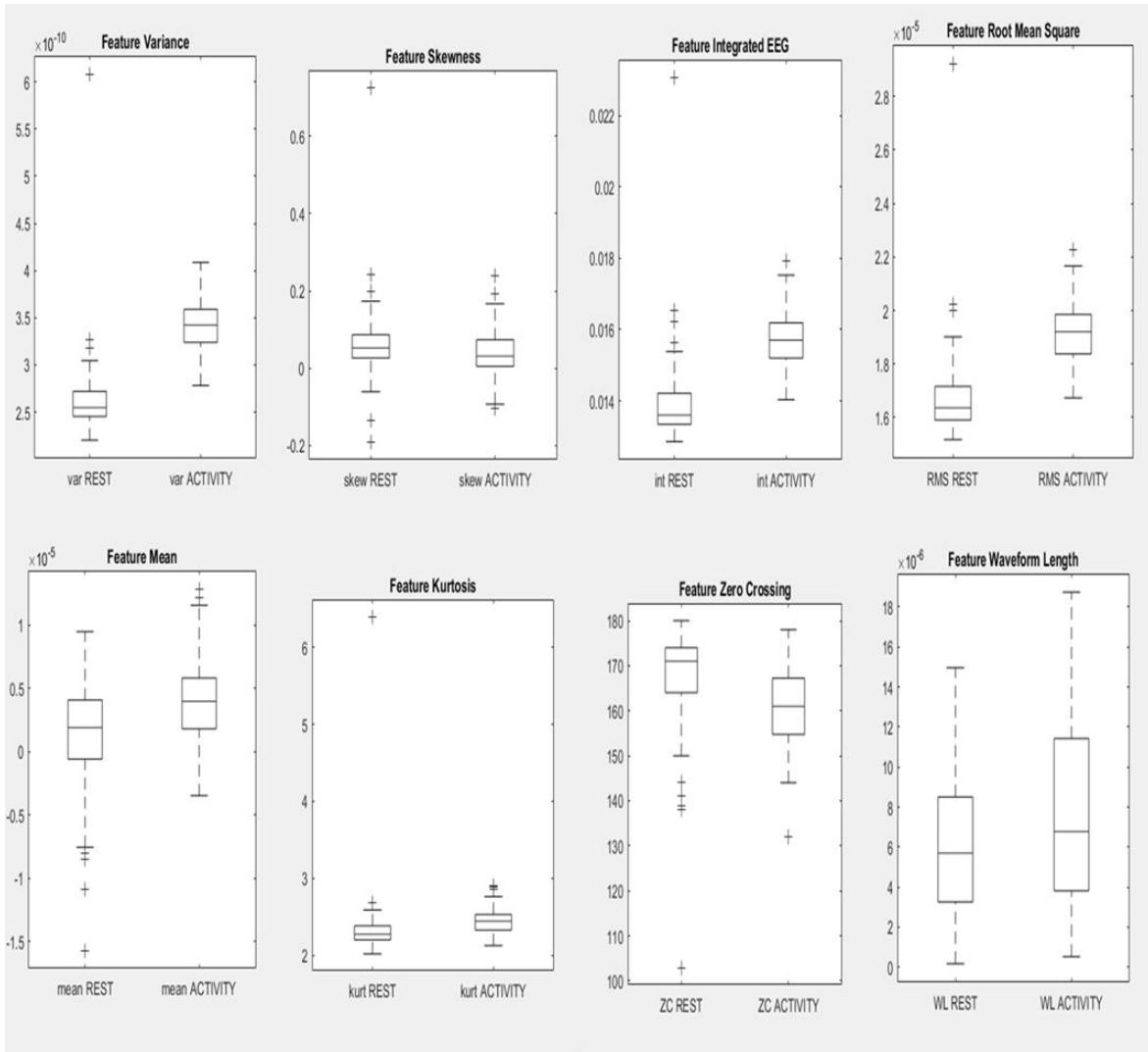
Figure 4.5. The Features Extracted from Time Series EEG Data of Resting and Activity States (Each Group of Boxes has 120 Values) (Data has been Taken from the Participant-One Day-One Channel CZ).

These boxplots show that some features may be useful to differentiate these two conditions. We investigated these features for all the participants in order to evaluate their successes and find the best features. After visual inspection of all features for all the

53

participants and the channels, the variance, root mean square and integrated EEG features were determined as useful features for differentiating resting and activity conditions (since they have similar results for all the participants). Figure 4.6 shows the boxplots for variance feature for all five participants. This feature was extracted from the channel CZ and each group of boxes includes 120 values.
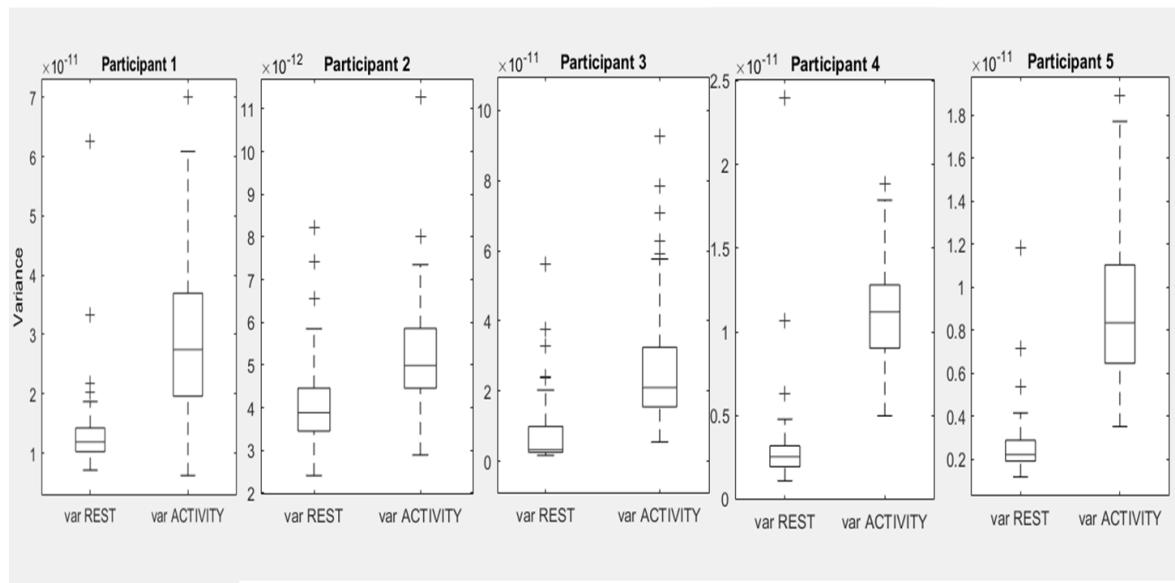
Figure 4.6. The Boxplots of Feature Variance from all Five Participants.

Figure 4.6 shows that the variance features of activity condition mostly have bigger values than the variance features of resting condition. This means that hand and finger movements cause some changes in variance of time series EEG signal. But we still can't declare that the changes in the variance features are caused by hand and finger artifacts, because the EEG data that we used so far was unfiltered EEG signal and the changes in the variance feature may be caused by the cognitive effects.

In order to illuminate this problem, the same features are extracted from filtered signals (10-25 Hz, 30-50 Hz and 50-100 Hz). In figure 4.7, boxplots of variance feature that is extracted from three filtered signals are shown.
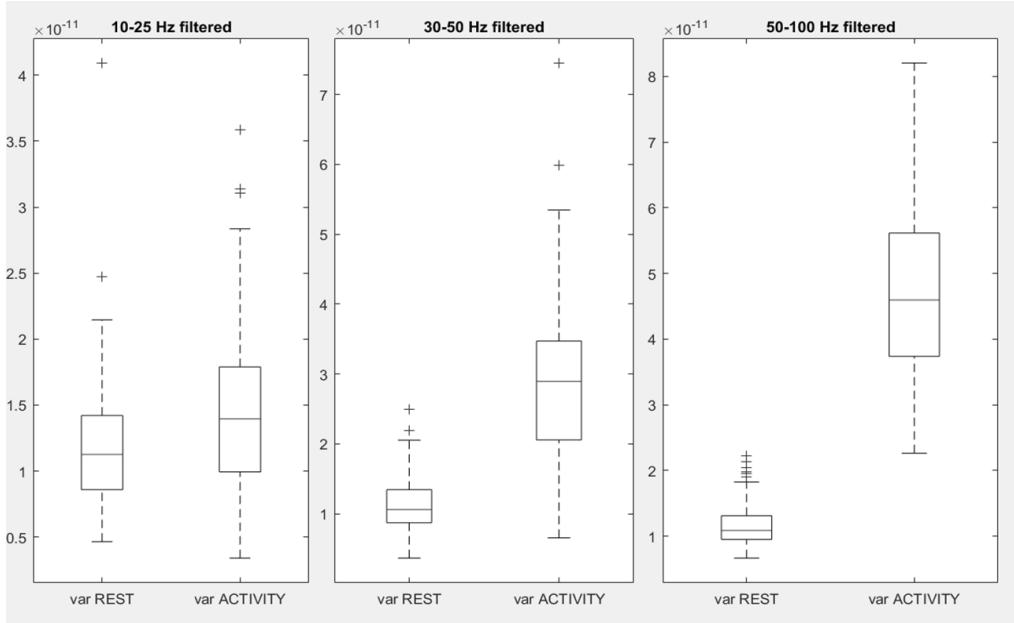
Figure 4.7. The Boxplots of Feature Variance from Three Filtered Signal

It can be observed from figure 4.7 that the difference between variances of two conditions becomes more visible in upper frequency bands. Since the upper frequency bands don't include much cognitive components, this result demonstrates the artefactual effects of hand and finger movements.

Since 50 to 100 Hz frequencies don't include much neural activity and 50-100 Hz filtered data yielded better results to differentiate two conditions, 50-100 Hz filtered data was used for the hand and finger artifact detection.

For this purpose, signals from all participants were filtered (50-100 Hz) and 8 different features have been extracted from the data of all participants for the resting and

the activity conditions. After that, each feature was grouped and compared. We investigated all the features for all the channels. As a result, we chose 3 features (variance, integrated EEG and RMS) from different channels in order to form the feature matrixes. And channels F3, T5, PZ and T6 have been selected by visual and statistical inspections. Figure 4.8, figure 4.9, and figure 4.10 show the boxplots of the selected features (variance, integrated EEG and Root Mean Square) respectively.
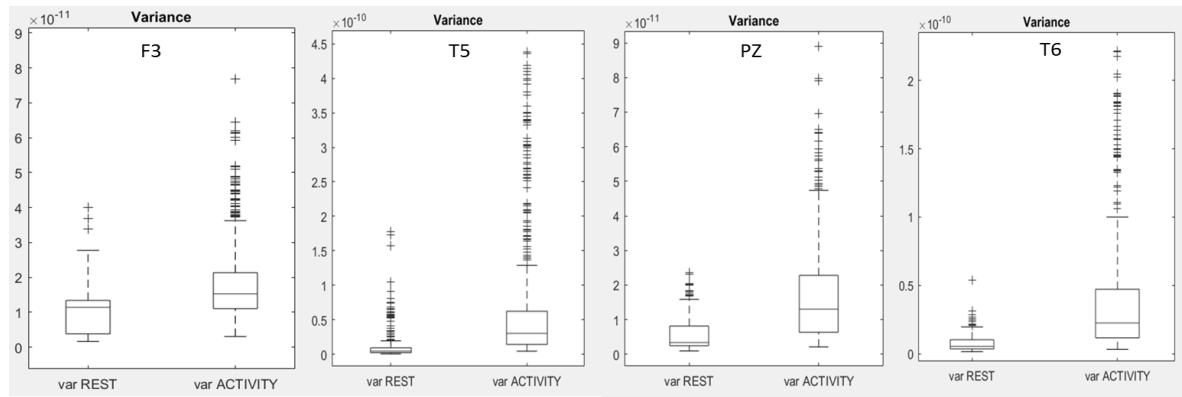

**Figure 4.8.** The Boxplots of Feature Variance (Channels F3, T5, PZ and T6).
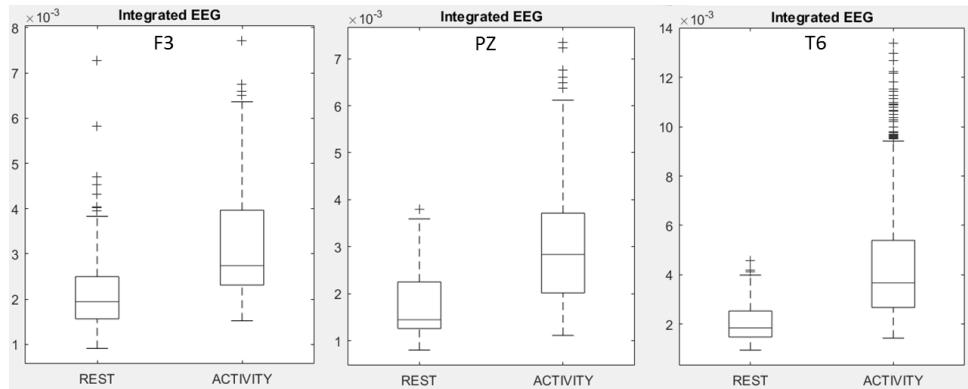
Figure 4.9. The Boxplots of Feature Integrated EEG (Channels F3, T5, PZ and T6).
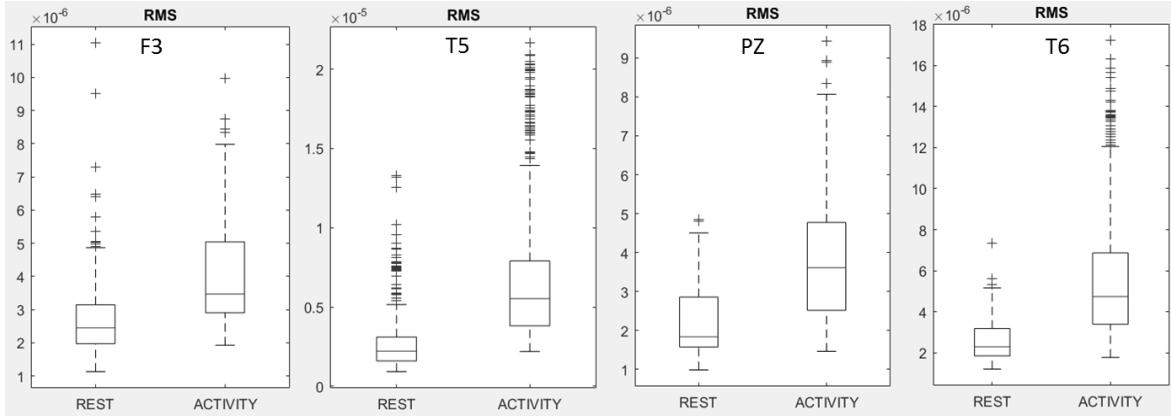
Figure 4.10. The Boxplots of Feature RMS (Channels F3, T5, PZ and T6).

In these figures we gathered the feature values of all 5 participants. One of the boxplots represents the feature values from the resting condition and the other one represents the values of the same feature that were extracted from the activity condition. By visual inspection, it can be deduced that the values of these features are mostly bigger for the activity condition.

After obtaining feature matrixes from the time series EEG data (50-100 Hz band pass filtered), the next step is to form neural network models for classification. It will be presented in section 4.4.

4.3.2. Right and Left Finger Keyboard Press Conditions

Resting and Activity conditions are easier to differentiate; since one of them is heavily contaminated by the small muscle movement effects (hand and finger artifacts). We were able to observe their effects using time series EEG signal. On the other hand, detecting the EEG portions that are affected by a single finger movement is a much more challenging problem. In order to present a useful study for Human Machine Team researchers, it is important to understand even the single finger movement effects.

In order to investigate the artefactual effects of single finger movements, we applied the same feature extraction process to the right and the left index finger keyboard press conditions. For this aim, as described in methodology part, we divided the data into 0.5-second chunks and separated them into two groups by using digital and EMG channels. One of the groups of chunks has no finger artifacts, while the other group includes finger artifacts. After dividing the EEG data (right finger press condition) into chunks, we had 900 chunks in total (639 chunks with no finger artifact and 261 chunks with finger artifact).

We extracted the time domain features (variance, RMS, mean, skewness, kurtosis and integrated EEG) from filtered (70-100 Hz) EEG data and compared two groups. But, these features didn't provide any consistent and decisive results as we got while comparing the rest and the activity states. This means that the artefactual effects of the finger movements can't be visually characterized just by using the time domain features. Because of that, another dimension of the signal was investigated. Time series EEG data has been converted into the frequency domain as described in methodology section. After this conversion, we extracted five frequency domain features (Average frequency power, maximum frequency power, frequency that has the max power, min frequency power and root mean square of the frequency powers). But, these frequency features didn't prove to be consistent and decisive either.

While the effects of the hand/finger artifacts can be observed by visual inspection in high contamination levels (rapid and continuous movements by both hands and fingers), we can't observe the effects by evaluating the features separately in low

contamination levels (single finger movement). We also applied the same process to the left finger keyboard press data and mouse movement data but the results were similar.

4.4. Detection of Hand and Finger Artifacts

Features extracted from time and frequency domain EEG data indicated that the finger and hand artifacts create observable effects on EEG signal. In this section, our study creates neural network models and trains these models by using feature matrixes that we formed in the previous section. After training these models, we test how these models perform in detection of the hand and finger artifacts.

4.4.1. Classification of Resting and Activity Conditions

Features extracted from filtered (50-100 Hz) time domain EEG data has been investigated for all the participants and channels and three features (variance, integrated EEG and RMS) and three channels (F3, PZ, T6) has been selected to form the feature matrixes. We keep the input numbers low, since these three features were good at separating two groups. We kept the classification model as simple as possible.

While forming the training and the test data, the first 35 seconds of the Resting and Activity data was used as training data, the last 15 seconds was used as test data and the 10 seconds of data between training and test data wasn't included to any group. We aimed to reduce the correlation between training and test data by using 10-second separation. Figure 4.11 shows this process visually.

REST/ ACTIVITY Data   60 Seconds

Training Data 35 secs	Separation 10 secs	Test Data 15 secs

Figure 4.11. Training and Test Data from the Rest and Activity Conditions.

After dividing rest and activity data of all participants into the training and test sets, chunks from resting condition was classed as "Zero" (0) and the chunks from activity condition was classed as "One" (1). Training and test sets were randomized after adding the class tags.

After creating training and test sets, we formed the neural network model. As mentioned in methodology chapter, neural network pattern recognition tool of MATLAB was used to form neural networks. Our model has 9 inputs and gives two outputs. We used the validation accuracy while determining the number of the hidden nodes. We first determined 5 for the number of the hidden layer nodes and the accuracy was 87.3% and the accuracy improved when we increased the hidden layer nodes. The accuracy improved to 92.2% when the hidden node number was 20. We continued to increase the number of the hidden nodes (we tried 30, 50 and 100), but there was definitive improvement on the accuracy. As a result, we decided to use 20 nodes for the hidden layer. Figure 4.12 demonstrates this neural network model.
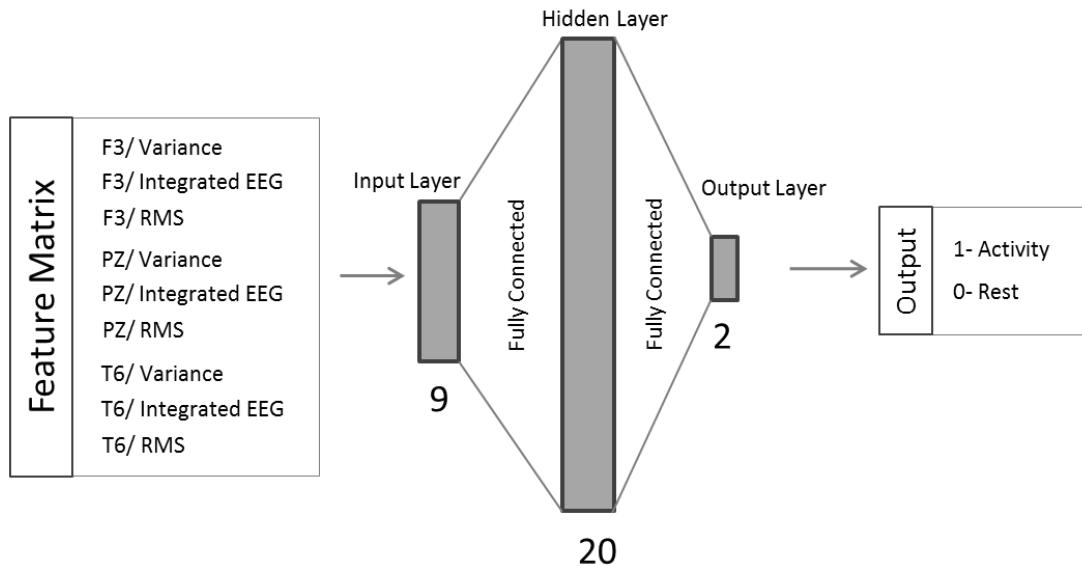
Figure 4.12. Neural Network Model for Classification of Activity and Resting Data

This model was trained by the training set and after the training, test data was evaluated on the trained model. The training set had 594 samples (feature matrix of 9x594) and the test set had 306 samples. Figure 4.13 presents the classification results of the test data by the confusion matrix.

	Total Population 306	**Predicted Condition**		
		Activity	Rest	
True Condition	Activity	134	5	96.4% 3.6%
	Rest	19	148	88.6% 11.4%
		87.6% 12.4%	96.7% 3.3%	92.20% 7.80%

Figure 4.13. Classification Results of the Rest and Activity Chunks (Test Data).

Figure 4.14 shows the ROC curve of this classification.
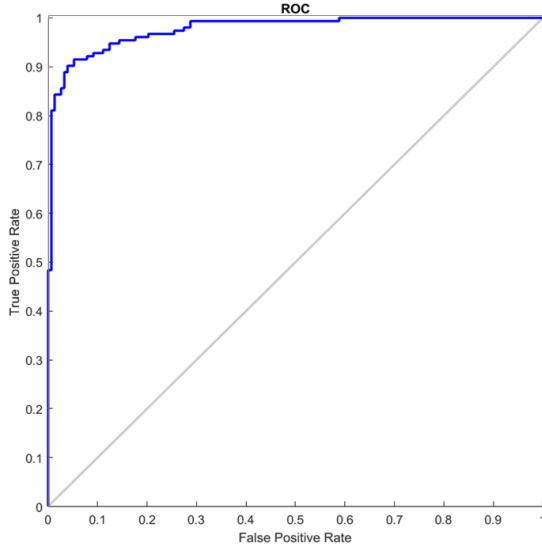
Figure 4.14. ROC Curve of the Test Data Classification

According to figure 4.13, the results showed that the chunks which belong to the activity data can be classified with the accuracy of 92.2%. The sensitivity (true positive rate, classifying the activity signal correctly) was 96.4% and the specificity (classifying the rest data correctly) was 88.6%.

It means that, 92.2% of the time, this model can classify the data correctly as resting data or activity data.

4.4.2. *Classification of Finger Artifacts*

Classification of the finger artifacts basically includes detecting the EEG segments that includes finger movements in it by using the artefactual effects of the finger movements. In this classification problem, the neural network model takes time and frequency domain features as the input and gives one of the two outputs (no finger

artifact or finger artifact). The time domain features were extracted from 70-100 Hz band passed filtered EEG data. In addition to this, we also band passed filtered the data between 30-50 Hz and extracted the same time domain features from this filtered data. This obtained us additional 6 features per channel. In addition to time domain features, 5 features were extracted from the frequency domain EEG data (70-100 Hz). As a result, we collected 17 features per channel. Figure 4.15 demonstrates these features.
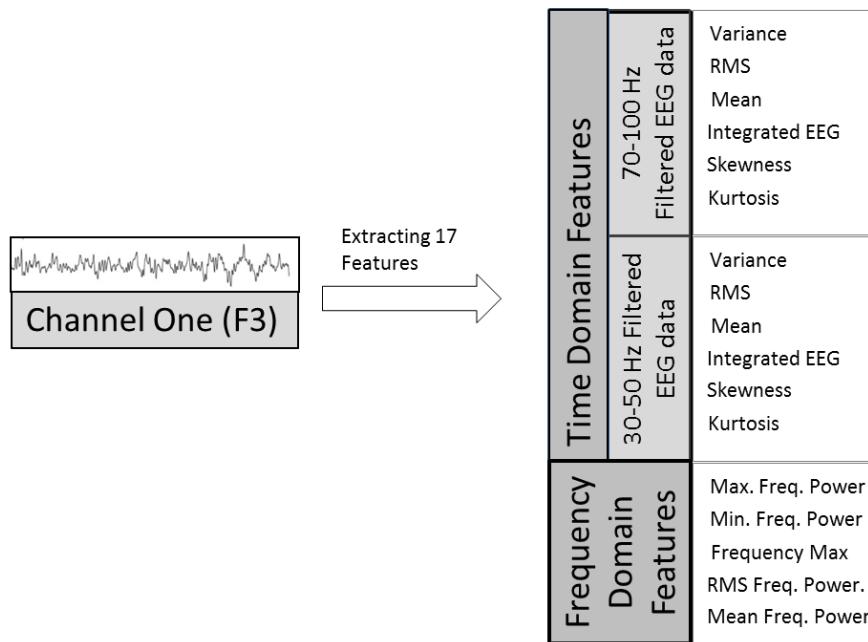

Figure 4.15. Extracted Features from one Channel

After extracting 17 features per channel, we obtained 153 (9x17) features. It means that, we extracted 153 features per chunk (0.5-second EEG data). We formed the feature matrixes by gathering these features from all chunks. And this feature matrix was used to train our model.

The neural network model had 153 nodes for the input layer. We decided the hidden layer node number as 200 and the output layer had two nodes (finger artifact or no artifact). Figure 4.16 demonstrates this neural network model.
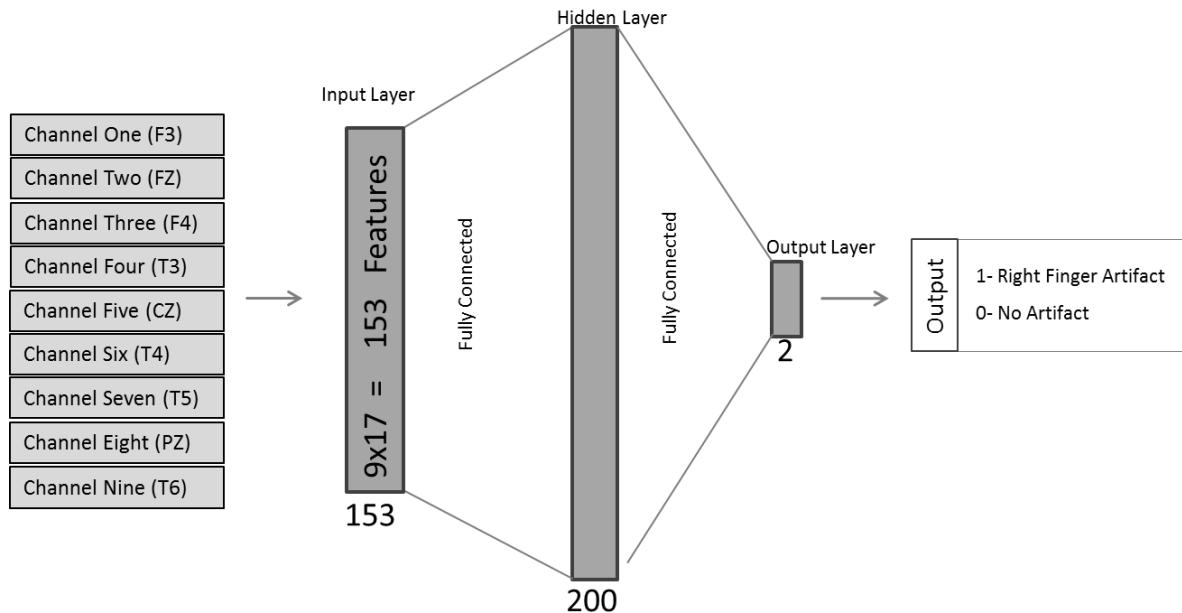
Figure 4.16. Neural Network Model for Classification of Right Finger Artifacts

As we mentioned earlier, for the right finger press conditions we had 900 chunks (634 of these chunks includes no finger artifact and 266 of them includes finger artifacts). We formed the feature matrix by using these 900 chunks (153x900-feature matrix). In order to create the training and the test sets, the feature matrix randomized and divided into two groups (70% training set and 30% test set).

We trained our model (shown in Figure 4.17) with the training set and test it with the test set. But the results were not good. The accuracy was 71.6%, but this value may

be misleading, because number of the groups was unbalanced (634 of these chunks includes no finger artifact and 266 of them includes finger artifacts). Because of that, we balanced the number of the groups as 266 chunks from each group. Training and tests sets were balanced and the model trained and tested again. The training data had 372 chunks and the test had 160 chunks. The test set had 80 chunks include finger artifact and 80 chunks that includes no artifact.

According to test results the accuracy was 65%. The results are shown by a confusion matrix in figure 4.17.

		Predicted Condition		
	Total Population 160	Finger Artifact	No Finger Artifact	
True Condition	Finger Artifact	43	19	69.4% 30.6%
	No Finger Artifact	37	62	62.6% 37.4%
		53.7% 46.3%	76.3% 23.7%	65.2% 34.8%

Figure 4.17. Classification Results of the Right Finger Artifacts.

According to figure 4.17, the results showed that the chunks which include right finger artifact and no artifact can be classified with the accuracy of 65%. The sensitivity (true positive rate, classifying the right finger artifact correctly) was 69.4% and the specificity (classifying the data that doesn't contain a finger artifact) was 62.6%.

This training and testing process has been made 10 times, in order to find an average accuracy performance. Table 4.1 shows 10 different test results for the classification model.

Table 4.1. Classification Results of Right Finger Artifact

Trainings	Specificity	Sensitivity	Accuracy
1	53.7	76.3	65.2
2	50	71.8	60.9
3	52.4	74.4	63.4
4	57.3	70.5	63.9
5	52	78.2	65.1
6	52.4	67.9	60.15
7	51.2	75.6	63.4
8	56.2	71.3	63.75
9	60.7	69.4	65.05
10	59.6	75.6	67.6

According to Table 4.1, the chunks which include the right finger artifact and no artifact can be classified with an average accuracy of 64%. This is not a good result for a classification and it doesn't provide enough evidence about the artefactual effects of the finger movements. Because of that, we tried to find ways to improve this classification accuracy.

Ocular artifact affects the EEG signal as we stated in the Literature Review Section. But in our model, we didn't consider these artifacts. Because of that, we decided to find and remove chunks that include eye blink artifacts and wrote a code that finds these chunks. This code uses the vertical EOG channel to determine when the participant blinks.

It was found that, 72 of the chunks (900 chunks in total) had eye blink artifacts and 8 of the chunks had both eye blink and finger artifacts. We removed these chunks from the feature matrix. After this process, we had 258 chunks that include finger artifacts and 562 chunks that include no artifact. We formed a balanced feature matrix by

taking 258 chunks from each group (516 Chunks in total) and 160 of these chunks (80 chunks from each group) were used as test data.

As a result, we created the same conditions with the previous classification (the classification that yielded 64% average accuracy). We only removed the chunks that include eye blink artifacts). We trained our model again with the new training set. We finally tested the trained model by using the test set. The results are shown by a confusion matrix in figure 4.18.

		Predicted Condition		
	Total Population 160	Finger Artifact	No Finger Artifact	
True Condition	Finger Artifact	57	24	70.4% 29.6%
	No Finger Artifact	19	60	75.9% 24.1%
		75% 25%	71.4% 28.6%	73.1% 26.9%

Figure 4.18. Classification Results of the Right Finger Artifacts (Blink Artifacts Removed).

According to figure 4.18, the classification accuracy improved to 73.1%. This training and testing process has been made 10 times, in order to find an average accuracy performance. Table 4.2 shows 10 different test results for the classification model.

Table 4.2. Classification Results of Right Finger Artifacts (Blink Artifacts Removed)

Trainings	Specificity	Sensitivity	Accuracy
1	70.4	75.9	73.1
2	69.7	71.4	70.55
3	71.1	75	73.05
4	71.1	69	70.05
5	67.8	77.4	72.6

6	69.7	69	69.35
7	65.8	77.4	71.6
8	66.8	76.2	71.5
9	64.5	81	72.75
10	69.7	79.8	74.75

According the table 4.2, the chunks which include right finger artifact and no artifact can be classified with an average accuracy of 72%. It can be observed that the accuracy of the classification increased to 72% from 64% after we removed the chunks that include blink artifacts. In addition to this, the average sensitivity (true positive rate, classifying the right finger artifact correctly) increased to 69.1% from to 55.3%.

It means that 72% of the time our model can detect whether the data (0.5-second segment) contains finger artifact or not. The detection performance still poor, but we observed that we could improve the accuracy (the performance the finger artifact detection) by eliminating the other artifacts.

4.5. Summary

In this chapter, experimental results and analysis from the methods described in Chapter III were presented. The investigative questions along with the answers supported in this chapter are summarized below:

Q1. Do small muscle movements (finger and hand movements caused by keyboard and mouse usage) have observable effects on EEG signals? Frequency power density plots and feature comparisons showed that the effects of the small muscle movements are observable when the data is heavily contaminated by these small muscle movements (activity and resting conditions). If we are dealing with really small muscle

movements such as a single finger keyboard press, their effects on EEG signal are not observable.

Q2. Is it possible to detect small muscle movements (hand and finger movements) from EEG signals, without any reference channel indicating hand/finger movement?

Classification results showed that it is possible to detect small muscle movements with a simple neural network models if the contamination level is high (activity vs, resting states). Our classification model (activity and resting data classification) yielded 92.2% accuracy. On the other hand, if the contamination level is really low (data contains a single finger movement), it is not possible to detect these artifacts with simple models. In order to finger artifact classification, we extracted 153 features (3 features in the previous model) and eliminated eye artifacts and after those improvements, our classification model yielded 72% accuracy. This is not a high accuracy but, it gives the idea about the effects of finger artifacts on the EEG signal.

As a result, we analyzed the best case and the worst-case scenarios in order to investigate the small muscle movement artifacts. And the results showed that these small muscle artifacts have effects on the EEG signals.

V. Conclusion and Future Work

Human Machine Teams (HMTs) and Brain Computer Interface (BCI) Systems are strongly rely on the EEG signals. Both of them need to use pure EEG signal that represents the neural activity of the brain. But the physiological and non-physiological artifacts distort the EEG signals and make the interpretation of cognitive state harder or may cause misinterpretations.

This thesis study focuses on a kind of muscle artifact that is caused by the small muscle movements, since the artefactual effects of these movements haven't been investigated in the literature. The effects of these movements have been investigating by using the best and the worst-case scenarios. In the best-case scenario, the EEG data has been heavily contaminated by the small muscle movements (activity state). The participants continuously move their both hands and fingers in this case. In the worst-case scenario, the data only contains single finger artifacts (right/left finger keyboard presses). And we analyzed these cases in order to investigate the effects of the small muscle movements on the EEG signal.

5.1. Research Findings

This thesis study investigates artefactual effects of the small muscle movements. For this aim, we investigate the best and the worst-case scenarios. For the best case (activity versus resting state) scenario, the results showed that the effects of the small muscle artifacts can be visually observed by extracting and comparing time domain features (such as variance and RMS) from the EEG signal. In this condition, the detection of these artifacts can be made by a simple neural network model with high accuracy. In

our study, the model we created can detect the data segments that contain small muscle artifacts with the accuracy of 92.2%. We used just the time domain of the EEG signal for this detection model. The accuracy of the model may be increased by adding some features from the frequency domain of the signal. But we didn't make further inspection in this scenario, because this accuracy level was enough to state that if the EEG data is heavily contaminated by small muscle artifacts, the effects can be detected.

For the worst-case scenarios (right/left finger presses), the results were different. We extracted the time and frequency domain features from the EEG signals in order to find some features that may visually show the effects of the finger artifacts. But none of the features provided visual evidence about finger artifacts. Because of this, we created neural network models and made classification to detect the finger artifacts. The results showed that the detection accuracy was 64%. This accuracy level was not enough to state that the finger artifacts are detectable. We applied the same process after eliminating the eye blink artifacts and the accuracy of the classification model increased to 72%.

5.2. Future Research

In this study, our model made the finger artifact detection with 64% accuracy. We improved the accuracy of the finger detection to 72% by eliminating the eye blink artifacts. This accuracy level shows that the finger movements have some effects on the EEG signal and we can improve their detection performance by cleaning other artifacts. It is obvious that the finger movements have really small effects on EEG signal when compared to large muscle movements (such as head movements) or the ocular artifacts. When we removed the eye blink artifact we got a significant improvement on the

detection. For the future work, other artifacts such as horizontal eye movements, other muscle movements may also be removed before investigating the finger artifacts. This may improve the accuracy of the detection. Since we have just 5 participants, our data was not enough to remove additional artifacts. In addition to this, we only recorded the EMG activity from the right arm and didn't capture other muscle activities.

In order to form the training and test datasets, we used the first parts of the recordings for training and the last parts of the data for testing sets. Since we had 5 participants, we didn't separate participants into two groups as training and test groups. For a future study, this experiment can be made by more participants and these participants can be divided into training and test groups. And by doing that, the test results may represent the success of the detection for common usage. In addition to this, the detection results may be tested with the resting and imaginary keyboard press conditions for a future study. When we test it with resting condition, we expect the model to classify the data portions as non-artefactual chunks, since these conditions include no finger artifacts. These results also demonstrate the success of the classifiers.

As a result, small muscle artifacts have detectable effects on the EEG signal. If the small muscle movements are intense, these effects can be observed visually. If the data has just finger movements, the effects cannot be observed visually but can be detected by machine learning methods such as neural networks.

Bibliography

Bozorgzadeh, Z., Birch, G. E., & Mason, S. G. (2000). The LF-ASD brain computer interface: on-line identification of imagined finger flexions in the spontaneous EEG of able-bodied subjects. IEEE. Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=859321

Fatourechi, M., Bashashati, A., Ward, R. K., & Birch, G. E. (2007). EMG and EOG artifacts in brain computer interface systems: A survey. *Clinical Neurophysiology*, *118*(3), 480–494. http://doi.org/10.1016/j.clinph.2006.10.019

Files, B. (2011). An introduction to EEG artifacts. *Perception*, (February), 1–67. http://doi.org/10.1016/B978-1-59749-281-2.00001-9

Gunaydin, O., & Ozkan, M. (2010). Design of a Brain Computer Interface System based on Electroencephalogram (EEG). *2010 4th European Education and Research Conference (EDERC)*, 150–154. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6151426

H. Klem, G., Otto Lüders, H., Jasper, H. H., & Elger, C. (1999). The ten twenty electrode system of the international federation. *Electroencephalography and Clinical Neurophysiology*, *10*(2), 370–375. http://doi.org/10.1016/0013-4694(58)90053-1

Klass, D. W. (1995). The Continuing Challenge of Artifacts in the EEG. *American Journal of EEG Technology*, *35*, 239–269.

Li, Y., Gao, X., Liu, H., & Gao, S. (2004). Classification of single-trial electroencephalogram during finger movement. *IEEE Transactions on Biomedical Engineering*, *51*(6), 1019–1025. http://doi.org/10.1109/TBME.2004.826688

Lisogurski, D., Birch, G. E. (1998). Identification of finger flexions from continuous EEG as a brain computer interface. Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society.(Vol. 4, pp. 2004–2007). Retrieved from http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=746997

Liu, Aiping; Chen, Xun; Chiang, Joyce; Wang, Z. Jane; McKeown, Martin; Ward, R. K. (2016). Removing muscle artifacts from EEG data: Multichannel or single-channel techniques? *To Appear, IEEE Sensors Journal*, *16*(7), 1986–1997.

NINDS. (2012). Brain Basics: Know your Brain. *National Institute of Neurological Disorders and Stroke*. Retrieved from http://www.ninds.nih.gov/disorders/brain_basics/know_your_brain.htm

Ochoa, J. B. (2002). *EEG Signal Classification for Brain Computer Interface Applications*. *Ecole Polytechnique Federale De Lausanne* (Vol. 7). http://doi.org/10.1.1.134.6148

Oja, E., & Hyvärinen, A. (2000). Independent Component Analysis: Algorithms and Applications, *1*(1).

O'Regan, S. H. (2013). Title Artefact detection and removal algorithms for EEG diagnostic systems. Retrieved from http://hdl.handle.net/10468/1391

Quinonez, D. (1998). Common applications of electrophysiology (EEG) in the past and today: The technologist's view. *Electroencephalography and Clinical Neurophysiology*, *106*(2), 108–112. http://doi.org/10.1016/S0013-4694(97)00112-0

Safieddine, D., Kachenoura, A., Albera, L., Birot, G., Karfoul, A., Pasnicu, A., … Merlet, I. (2012). Removal of muscle artifact from EEG data: Comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches. *EURASIP Journal on Advances in Signal Processing*, *2012*, 127: 1-15. http://doi.org/10.1186/1687-6180-2012-127

Sweeney, K. T., Ward, T. E., & McLoone, S. F. (2012). Artifact removal in physiological signals-practices and possibilities. *IEEE Transactions on Information Technology in Biomedicine*, *16*(3), 488–500. http://doi.org/10.1109/TITB.2012.2188536

Tan, D., & Nijholt, A. (2010). Brain-Computer Interfaces, 3–19. http://doi.org/10.1007/978-1-84996-272-8

Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement Science Review*, *2*, 1–11. http://doi.org/10.1021/pr070350l

Urigüen, J. A., & Garcia-Zapirain, B. (2015). EEG artifact removal-state-of-the-art and guidelines. *Journal of Neural Engineering*, *12*(3), 31001. http://doi.org/10.1088/1741-2560/12/3/031001

Vanhatalo, S., Voipio, J., Dewaraja, A., Holmes, M. D., & Miller, J. W. (2003). Topography and elimination of slow EEG responses related to tongue movements. *NeuroImage*, *20*(2), 1419–1423. http://doi.org/10.1016/S1053-8119(03)00392-6

Varghese, J. P. (2009). Analysis of EEG Signals For EEG-based Brain-Computer Interface, (July), 1–54.

Vigneshwari, C., Vimala, V., Vignesh, S. V., & Sumithra, G. (2013). Analysis of Finger Movements Using EEG Signal. *International Journal of Emerging Technology and Advanced Engineering*, *3*(1), 583–588.

Whitham, E. M., Pope, K. J., Fitzgibbon, S. P., Lewis, T., Clark, C. R., Loveless, S., … Willoughby, J. O. (2007). Scalp electrical recording during paralysis: Quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clinical Neurophysiology*, *118*(8), 1877–1888. http://doi.org/10.1016/j.clinph.2007.04.027

Zhang, C., Tong, L., Zeng, Y., Jiang, J., Bu, H., Yan, B., & Li, J. (2015). Automatic Artifact Removal from Electroencephalogram Data Based on A Priori Artifact Information. *BioMed Research International*, *2015*, 1–8. http://doi.org/10.1155/2015/720450

1. REPORT DATE *(DD-MM-YYYY)* 12-22-2016	2. REPORT TYPE Master's Thesis	3. DATES COVERED *(From – To)* Aug 2015 – Dec 2016
4. TITLE AND SUBTITLE Analysis of Small Muscle Movement Effects on EEG Signals		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) YANTERI, Erhan E., 1st Lt., TURAF		5d. PROJECT NUMBER 16G347
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Way, Building 640 WPAFB OH 45433-8865		8. PERFORMING ORGANIZATION REPORT NUMBER AFIT- ENG-MS-16-D-051
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) AFIT/ENR 2950 Hobson Way WPAFB OH 45433-7765		10. SPONSOR/MONITOR'S ACRONYM(S) AFOSR (managed by AFIT Faculty Research Council)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES
This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT

Developments in the biomedical signal processing have led the electroencephalography (EEG) to be a critical tool for the Brain Computer Interface (BCI) systems and Human Machine Teams (HMTs). Both of them strongly rely on the EEG signals in order to evaluate the neural activity and the cognitive state. But the physiological and non-physiological artifacts distort the EEG signals and make the interpretation of cognitive state harder or may cause misinterpretations. While interacting with computers, humans perform small motor muscle movements such as operating a keyboard and mouse. On the other side, the computer agent needs to know the cognitive state of the human teammate in order to make decisions and the EEG signals are the only information source of cognitive state.

In this thesis, the artefactual effects of the small muscle movements were investigated. Upper frequency bands (>30 Hz) of the EEG signal were extracted in order to investigate the artefactual effects of the small muscle movements. When the contamination level is high, the detection of the small muscle artifact can be made with the 92.2% accuracy. If these artifacts are really small such as a single finger movement, the detection accuracy decreases to 64%. But, the detection accuracy increases to 72% after removing the eye blink artifacts. The results of the classification support our hypothesis about the artefactual effects of the small muscle movements.

15. SUBJECT TERMS
EEG, Artifact Detection, Small Muscle Movements, Human Machine Teams, Machine learning

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Dr. Brett Borghetti, AFIT/ENG
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*
U	U	U	UU	90	(937) 785-6565, ext. 4612 brett.borghetti@afit.edu